

THE ACCOUNTING REVIEW

VOLUME 86, No. 3 MAY 2011

CHRISTOPHER P. AGOLIA, TIMOTHY S. DOUPNIK, and GEORGE T. TSAKUMIS

Principles-Based versus Rules-Based Accounting Standards: The Influence of Standard Precision and Audit Committee Strength on Financial Reporting Decisions

ANWER S. AHMED, EMRE KILIC, and GERALD J. LOBO

Effects of SFAS 133 on the Risk Relevance of Accounting Measures of Banks' Derivative Exposures

ANIL ARYA and BRIAN MITTENDORF

Supply Chains and Segment Profitability: How Input Pricing Creates a Latent Cross-Segment Subsidy

JEAN C. BEDARD and LYNFORD GRAHAM

Detection and Severity Classifications of Sarbanes-Oxley Section 404 Internal Control Deficiencies

JEREMY BERTOMEU, ANNE BEYER, and RONALD A. DYE

Capital Structure, Cost of Capital, and Voluntary Disclosures

JENNIFER L. BLOUIN, JANA S. RAEDY, and DOUGLAS A. SHACKELFORD

Dividends, Share Repurchases, and Tax Clientele: Evidence from the 2003 Reductions in Shareholder Taxes

BILLY E. BREWSTER

How a Systems Perspective Improves Knowledge Acquisition and Performance in Analytical Procedures

DAIN C. DONELSON, ROSS JENNINGS, and JOHN McINNIS

Changes over Time in the Revenue-Expense Relation: Accounting or Economics?

BENG WEE GOH and DAN LI

Internal Controls and Conditional Conservatism

GUOJIN GONG, LAURA YUE LI, and JAE YONG SHIN

Relative Performance Evaluation and Related Peer Groups in Executive Compensation Contracts

YANIV KONCHITCHKI

Inflation and Nominal Financial Reporting: Implications for Performance and Stock Prices

REUVEN LEHAVY, FENG LI, and KENNETH MERKLEY

The Effect of Annual Report Readability on Analyst Following and the Properties of Their Earnings Forecasts

A JOURNAL OF THE
AMERICAN ACCOUNTING ASSOCIATION

AMERICAN ACCOUNTING ASSOCIATION

The members of the 2010–2011 Executive Committee of the Association are as follows:

President	Kevin D. Stocks, Brigham Young University
President-Elect	Gregory B. Waymire, Emory University
Past President	Nancy A. Bagranoff, Old Dominion University
Vice President	Ira Solomon, University of Illinois at Urbana–Champaign
Vice President–Education	Belverd E. Needles, Jr., DePaul University
Vice President–Education–Elect	Alan Reinstein, Wayne State University
Vice President–Finance	Stacy E. Kovar, Kansas State University
Vice President–International	Kazuo Hiramatsu, Kwansei Gakuin University
Vice President–Professional Relations	Robert H. Colson, Grant Thornton, LLP
Vice President–Publications	Jean C. Bedard, Bentley University
Vice President–Publications–Elect	Christine A. Botosan, The University of Utah
Vice President–Research	John Christensen, University of Southern Denmark
Vice President–Sections and Regions	Gail Hoover King, Rockhurst University

The American Accounting Association was founded in 1916 as the American Association of University Instructors in Accounting. Its present name was adopted in 1936. The Association is a voluntary organization of persons interested in accounting education, research, and practice.

Membership in the Association is available at the following annual rates:

Full membership:

The membership dues include *Accounting Education News*, and online selection of one, two, or all three of the following Association-wide journals: *The Accounting Review*, *Accounting Horizons*, and *Issues in Accounting Education*. (This does not include Section journals.) Select from:

<i>The Accounting Review</i> only	U.S. \$185.00	with Print \$220.00
<i>Accounting Horizons</i> OR <i>Issues in Accounting Education</i> only	U.S. \$185.00	with Print \$200.00
<i>The Accounting Review</i> AND <i>Accounting Horizons</i> OR <i>Issues in Accounting Education</i>	U.S. \$195.00	with Print \$245.00
<i>Accounting Horizons</i> AND <i>Issues in Accounting Education</i>	U.S. \$195.00	with Print \$225.00
All three Association-wide journals	U.S. \$205.00	with Print \$270.00

Associate membership:

Full-time students are eligible for associate membership. The print option is not available for Associate members.

Select from:

Any one Association-wide journal	U.S. \$25.00
Any two Association-wide journals	U.S. \$35.00
All three Association-wide journals	U.S. \$45.00

In addition, full members may belong to one or more Sections: Accounting, Behavior and Organizations; Accounting Programs Leadership Group; American Taxation Association; Auditing; Diversity; Financial Accounting and Reporting; Forensic and Investigative Accounting; Gender Issues and Worklife Balance; Government and Non-profit; Information Systems; International Accounting; Management Accounting; Strategic and Emerging Technologies; Teaching, Learning and Curriculum; and Two-Year College. Each Section assesses its own annual dues, which are collected by the Association.

Through May 31, 2011, all editorial correspondence and manuscripts should be emailed to tar@mcombs.utexas.edu. Phone inquiries can be directed to Editorial Assistant Mary Capps at (512) 471-5318; Fax (512) 471-3904. Beginning June 1, 2011, please check online at <https://aaahq.org/AAAforms/journals/tarsubmit.cfm> for instructions on how to submit a new manuscript to the incoming editor, John Harry Evans III. Revisions of manuscripts originally submitted before June 1, 2011 should be sent via email to the incoming editor, John Harry Evans III, at tar@katz.pitt.edu.

The submission fee of \$200.00 for members of the AAA or \$400.00 for nonmembers should be paid online (VISA or MasterCard only) at: <https://aaahq.org/AAAforms/journals/tarsubmit.cfm> and must be paid when article is submitted. Request for reprint permission, subscriptions, advertisements, and dues should be sent to American Accounting Association, 5717 Bessie Drive, Sarasota, FL 34233-2399 • Telephone: (941) 921-7747 • Fax: (941) 923-4093 • Email: Info@aaahq.org. Notice of change of address should be mailed at least eight weeks in advance to ensure undelayed receipt of *The Accounting Review*.

The Accounting Review is indexed in the Social Science Citation Index which is included in Web of Science.

The Accounting Review (ISSN 0001-4826 print and ISSN 1558-7967 online) is published six times a year in January, March, May, July, September, and November by the American Accounting Association, 5717 Bessie Drive, Sarasota, Florida 34233-2399. General subscriptions are available at the annual rates of \$395 for a print copy or at \$445 for a print and electronic copy. The member subscription rate for a hard copy is \$35. Electronic access is included with membership dues. Periodical postage paid at Sarasota, Florida and additional mailing offices. **POSTMASTER:** Send address change to *The Accounting Review*, 5717 Bessie Drive, Sarasota, Florida 34233-2399.

THE ACCOUNTING REVIEW

A JOURNAL OF THE
AMERICAN ACCOUNTING ASSOCIATION

Senior Editor

Steven J. Kachelmeier

Editors

Shannon W. Anderson	D. Paul Newman
J. Harry Evans III	Thomas C. Omer
James E. Hunton	Wayne B. Thomas
Kathryn Kadous	Mark A. Trombley
Sanjay G. Kallapur	Beverly R. Walther
Ranjani Krishnan	Paul A. Zarowin
Laureen A. Malnes	

58

VOL. 86

MAY 2011

No. 3

CHRISTOPHER P. AGOLIA, TIMOTHY S. DOUPNIK, and GEORGE T. TSAKUMIS	
Principles-Based versus Rules-Based Accounting Standards: The Influence of Standard Precision and Audit Committee Strength on Financial Reporting Decisions	747
ANWER S. AHMED, EMRE KILIC, and GERALD J. LOBO	
Effects of SFAS 133 on the Risk Relevance of Accounting Measures of Banks' Derivative Exposures	769
ANIL ARYA and BRIAN MITTENDORF	
Supply Chains and Segment Profitability: How Input Pricing Creates a Latent Cross-Segment Subsidy	805
JEAN C. BEDARD and LYNFORD GRAHAM	
Detection and Severity Classifications of Sarbanes-Oxley Section 404 Internal Control Deficiencies	825
JEREMY BERTOMEU, ANNE BEYER, and RONALD A. DYE	
Capital Structure, Cost of Capital, and Voluntary Disclosures	857
JENNIFER L. BLOUIN, JANA S. RAEDY, and DOUGLAS A. SHACKELFORD	
Dividends, Share Repurchases, and Tax Clientele: Evidence from the 2003 Reductions in Shareholder Taxes	887
BILLY E. BREWSTER	
How a Systems Perspective Improves Knowledge Acquisition and Performance in Analytical Procedures	915
DAIN C. DONELSON, ROSS JENNINGS, and JOHN McINNIS	
Changes over Time in the Revenue-Expense Relation: Accounting or Economics?	945
BENG WEE GOH and DAN LI	
Internal Controls and Conditional Conservatism	975
GUOJIN GONG, LAURA YUE LI, and JAE YONG SHIN	
Relative Performance Evaluation and Related Peer Groups in Executive Compensation Contracts	1007
YANIV KONCHITCHKI	
Inflation and Nominal Financial Reporting: Implications for Performance and Stock Prices	1045
REUVEN LEHAVY, FENG LI, and KENNETH MERKLEY	
The Effect of Annual Report Readability on Analyst Following and the Properties of Their Earnings Forecasts	1087

BOOK REVIEWS, Stephen A. Zeff, Editor

Joan Luft and Michael D. Shields	
<i>Psychology Models of Management Accounting</i>	THERESA LIBBY 1117
Chris Poullaos and Suki Sian (editors)	
<i>Accountancy and Empire: The British Legacy of Professional Organization</i>	IGNACE DE BEELDE 1119
Capsule Commentary	STEPHEN A. ZEFF 1122
Editorial Policy and Style Information	1123

Editorial Staff of *The Accounting Review*
2010–2011

SENIOR EDITOR

Steven J. Kachelmeier, The University of Texas at Austin

EDITORS

Shannon W. Anderson, University of California, Davis

J. Harry Evans III, University of Pittsburgh

James E. Hunton, Bentley University

Kathryn Kadous, Emory University

Sanjay G. Kallapur, Indian School of Business

Ranjani Krishnan, Michigan State University

Laureen A. Maines, Indiana University

D. Paul Newman, The University of Texas at Austin

Thomas C. Omer, Texas A&M University

Wayne B. Thomas, University of Oklahoma

Mark A. Trombley, The University of Arizona

Beverly R. Walther, Northwestern University

Paul A. Zarowin, New York University

EDITORIAL ASSISTANT

Mary C. Capps, The University of Texas at Austin

RESEARCH ASSISTANT

Tracie Majors, The University of Texas at Austin

American Accounting Association

EXECUTIVE DIRECTOR

Tracey E. Sutherland

COMMUNICATIONS DEPARTMENT

Beverly Harrelson

Diane Hazard

Lisa Habbitz

Lindsay Wharton Peters

Kathy Casper

Judy Cothorn

Jan Kovarik (subcontracted)

Copyediting Team (subcontracted)

OTHER AAA STAFF

Dierdre Harris

David Boynton

Alex Metz

Pat Stein

Debora Gardner

Peggy Turczyn

Barbara Gutierrez

Barbara Brady

Arlene Oliver

Roland LaTulip

Diane Leger

Man Fong

Michelle Bowles

Kurt Gardner

Gay Williams

Editorial Advisory and Review Board

Anwer S. Ahmed, Texas A&M University
Ashiq Ali, The University of Texas at Dallas
Anil Arya, The Ohio State University
Benjamin C. Ayers, University of Georgia
William R. Baber, Georgetown University
Stephen P. Baginski, University of Georgia
Ramji Balakrishnan, The University of Iowa
Tim Baldenius, Columbia University
Linda Smith Bamber, University of Georgia
Orle E. Barron, The Pennsylvania State University
Anne L. Beatty, The Ohio State University
Jean C. Bedard, Bentley University
Robert J. Bloomfield, Cornell University
Jennifer L. Blouin, University of Pennsylvania
Mark T. Bradshaw, Boston College
Lawrence D. Brown, Georgia State University
Brian J. Bushee, University of Pennsylvania
Jeffrey L. Callen, University of Toronto
Joseph V. Carcello, The University of Tennessee
Mary Ellen Carter, Boston College
Paul K. Chaney, Vanderbilt University
Qiang Cheng, University of Wisconsin–Madison
Michael B. Clement, The University of Texas at Austin
William M. Cready, The University of Texas at Dallas
Antonio Davila, IESE Business School–University of Navarra
Patricia M. Dechow, University of California, Berkeley
Mark L. DeFond, University of Southern California
Hemang Desai, Southern Methodist University
Sunil Dutta, University of California, Berkeley
Leslie G. Eldenburg, The University of Arizona

Patricia M. Fairfield, Georgetown University
David B. Farber, University of Missouri–Columbia
Paul E. Fischer, The Pennsylvania State University
Jennifer Francis, Duke University
Jere R. Francis, University of Missouri–Columbia
James R. Frederickson, Melbourne Business School
Frank B. Giger, University of Minnesota
Cristi A. Gleason, The University of Iowa
Johnathan C. Glover, Carnegie Mellon University
Steven M. Glover, Brigham Young University
David A. Guenther, University of Oregon
Luzi Hall, University of Pennsylvania
Jeffrey Hales, Georgia Institute of Technology
John R. M. Hand, The University of North Carolina at Chapel Hill
R. Lynn Hannan, Georgia State University
Gilles Hilary, INSEAD
D. Eric Hirst, The University of Texas at Austin
Frank D. Hodge, University of Washington
Chris E. Hogan, Michigan State University
Ole-Kristian Hope, University of Toronto
Richard W. Houston, The University of Alabama
Raffi J. Indjejikian, University of Michigan
Christopher D. Ittner, University of Pennsylvania
Debra C. Jeter, Vanderbilt University
Karla M. Johnstone, University of Wisconsin–Madison
Bin Ke, Nanyang Technological University
S. Jane Kennedy, University of Washington
Inder K. Khurana, University of Missouri–Columbia
Michael D. Kimbrough, University of Maryland
Ronald R. King, Washington University in St. Louis
April Klein, New York University

W. Robert Knechel, University of Florida
Lisa L. Koonce, The University of Texas at Austin
Susan D. Krlische, American University
Jayanthi Krishnan, Temple University
William N. Lanen, University of Michigan
David F. Larcker, Stanford University
Reuven Lehavry, University of Michigan
Carolyn B. Levine, Carnegie Mellon University
Robert Libby, Cornell University
Joan L. Luft, Michigan State University
Anne M. Magro, George Mason University
Carol Marquardt, Baruch College-CUNY
Dawn A. Matsumoto, University of Washington
Edward Maydew, The University of North Carolina at Chapel Hill
Brian W. Mayhew, University of Wisconsin-Madison
Sarah E. McVay, University of Utah
Kenneth A. Merchant, University of Southern California
Gregory S. Miller, University of Michigan
Lillian F. Mills, The University of Texas at Austin
Brian Mittendorf, The Ohio State University
Partha S. Mohanram, University of Toronto
Steven J. Monahan, INSEAD
Donald V. Moser, University of Pittsburgh
James N. Myers, University of Arkansas
Venky Nagar, University of Michigan
Terry L. Neal, The University of Tennessee
Karen K. Nelson, Rice University
Edward F. O'Donnell, Southern Illinois University Carbondale
Per M. Olsson, Duke University
Chul W. Park, The University of Hong Kong
Mark E. Peecher, University of Illinois at Urbana-Champaign
Sean A. Pfeffer, University of Kentucky
Kathy Ruby Petroni, Michigan State University

John D. Phillips, University of Connecticut
Morton Pincus, University of California, Irvine
Joseph D. Piotroski, Stanford University
Marlene A. Plumlee, University of Utah
Suresh Radhakrishnan, The University of Texas at Dallas
Kannan Raghunandan, Florida International University
Shiva Rajgopal, Emory University
Scott Richardson, BlackRock Inc.
John R. Robinson, The University of Texas at Austin
Darren T. Roulstone, The Ohio State University
Katherine Schipper, Duke University
Karen L. Sedatole, Michigan State University
Catherine M. Shakespeare, University of Michigan
Lakshmanan Shivakumar, London Business School
K. (Shiva) Sivaramakrishnan, University of Houston
Hollis A. Skaffe, University of Wisconsin-Madison
Richard G. Sloan, University of California, Berkeley
J. Reed Smith, Indiana University at Indianapolis
Mark T. Soliman, University of Washington
Geoffrey B. Sprinkle, Indiana University
Mary Harris Stanford, Texas Christian University
Phillip C. Stocken, Dartmouth College
K. R. Subramanyam, University of Southern California
Hum-Tong Tan, Nanyang Technological University
Kristy L. Towry, Emory University
Senyo Tse, Texas A&M University
Jennifer Wu Tucker, University of Florida
Irem Tuna, London Business School
Mohan Venkatachalam, Duke University
Joseph P. Weber, Massachusetts Institute of Technology
Michael S. Wilkins, Texas A&M University
Richard H. Willis, Vanderbilt University
Arnold M. Wright, Northeastern University
Teri Lombardi Yohn, Indiana University
Xiao-Jun Zhang, University of California, Berkeley
Mark F. Zimbelman, Brigham Young University

PAST EDITORS

1926–1929	William A. Paton	1976–1978	Don T. DeCoster
1929–1943	Eric L. Kohler	1978–1983	Stephen A. Zeff
1944–1947	A. C. Littleton	1983–1987	Gary L. Sundem
1948–1949	Robert L. Dixon	1987–1990	William Kinney
1950–1959	Frank P. Smith	1990–1994	A. Rashad Abdel-khalik
1959–1962	Robert K. Mautz	1994–1997	Robert P. Magee
1962–1965	Lawrence L. Vance	1997–2000	Gerald L. Salamon
1965–1967	Wendell Trumbull	2000–2002	Linda Smith Bamber
1968–1970	Charles H. Griffin	2002–2005	Terry Shevlin
1971–1972	Eldon S. Hendriksen	2005–2008	Dan S. Dhaliwal
1973–1975	Thomas F. Keller		

Principles-Based versus Rules-Based Accounting Standards: The Influence of Standard Precision and Audit Committee Strength on Financial Reporting Decisions

Christopher P. Agoglia

University of Massachusetts Amherst

Timothy S. Douppnik

University of South Carolina

George T. Tsakumis

Drexel University

ABSTRACT: Recent accounting scandals have resulted in regulatory initiatives designed to strengthen audit committee oversight of corporate financial reporting and have led to a concern that U.S. GAAP has become too rules-based. We examine issues related to these initiatives using two experiments. CFOs in our experiments exhibit more agreement and are less likely to report aggressively under a less precise (more principles-based) standard than under a more precise (more rules-based) standard. Our results also indicate that CFOs applying a more precise standard are less likely to report aggressively in the presence of a strong audit committee than a weak audit committee. We find no effect of audit committee strength when the standard is less precise. Finally, we find support for a three-path mediating model examining mechanisms driving the effect of standard precision on aggressive reporting decisions. These results should be of interest to U.S. policymakers as they continue to contemplate a shift to more principles-based accounting standards (e.g., IFRS).

Keywords: *standard precision; rules-based standards; principles-based standards; audit committee; IFRS.*

Data Availability: *Contact the authors.*

We thank Wendy Bailey, Cathy Beaudoin, Bill Dilla, Rick Hatfield, Scott Jackson, Tamara Lambert, Justin Leiby, Steve Perreault, Roger Silvers, Hun-Tong Tan, Brad Tuttle, Xiaochuan Zheng, Lauren Maines (editor), Steve Kachelmeier (senior editor), two anonymous reviewers, and participants at the 2009 AAA Annual Meeting, the 2009 AAA International Section Midyear Conference, and the 2010 Deloitte/University of Kansas Auditing Symposium for their helpful comments and assistance.

Editor's note: Accepted by Lauren Maines.

*Submitted: April 2009
Accepted: December 2010
Published Online: April 2011*

I. INTRODUCTION

Our study examines the effect of accounting standard precision on financial statement preparers' reporting judgments, as well as the potential role that the audit committee plays in mitigating aggressive financial reporting under differing levels of standard precision.¹ A wave of corporate accounting scandals in recent years has led to a push for regulatory changes. As part of the Congressional plan to reform U.S. financial reporting, the Sarbanes-Oxley Act of 2002 (SOX, U.S. House of Representatives 2002) introduced major regulatory initiatives in an attempt to overhaul financial reporting and corporate governance systems. These initiatives include two remedies aimed at dampening aggressive financial reporting: (1) the potential adoption of principles-based accounting standards and (2) enhancement of public company audit committees.

Concern has been rising within the financial/investing community that U.S. accounting standards have become too "rules-based." With bright-line tests and detailed guidance, U.S. standards have become so precise that many feel they invite opportunistic interpretation by corporate executives. The perception that a significant number of executives have been concerned with meeting the letter of a rule, more so than its spirit, has led to a call to consider a more "principles-based" regime. Accordingly, SOX required the Securities and Exchange Commission (SEC) to conduct a study on the possible adoption of principles-based standards by the U.S. financial reporting system. Recently, the SEC moved the principles-versus-rules debate to the front burner by proposing a roadmap that could lead to the mandatory adoption of the more principles-based International Financial Reporting Standards (IFRS) by U.S. public companies (SEC 2008).²

Additionally, the Financial Accounting Standards Board (FASB) issued a proposal for a principles-based approach to U.S. standard-setting and asked interested parties to comment (FASB 2002). Many who responded were supportive of a principles-based model, believing that such an approach would lead to higher quality financial reporting with less opportunity to "exploit the gaps in GAAP" (CalPERS) or use "financial accounting engineering" to get around detailed, rules-based standards (PricewaterhouseCoopers).³ Conversely, some argued that less specificity in accounting standards would result in an increase in manipulation of financial results (e.g., Intel and former FASB member David Mosso). On the separate issue of inter-firm comparability, a number of letter-writers expressed concern that, with an increased reliance on judgment, implementation of principles-based standards could result in a decrease in comparability across firms (e.g., IBM, Pfizer, Goldman Sachs, BDO Seidman). The ongoing debate over whether and when to move toward more principles-based accounting standards reflects the uncertainty of both the anticipated desirable and undesirable effects of such a paradigm shift. Although these are empirical questions that can be addressed experimentally to inform policy makers prior to making a move to principles-based standards, the academic literature is limited with respect to research regarding these uncertainties (Maines et al. 2003).

As noted, the possible adoption of principles-based financial reporting standards is only one of the regulatory remedies embodied in SOX to improve U.S. financial reporting. Section 301 of SOX also expands the responsibilities of the audit committee and places greater focus on its role in enhancing the quality of financial reporting. Section 407 requires public companies to disclose

¹ Similar to Hackenbrack and Nelson (1996), we define aggressive financial reporting as the favorable portrayal of a company's financial situation even when that reporting is not clearly indicated by the facts.

² On February 24, 2010, the SEC issued a statement reaffirming its support for global accounting standards while pushing back the earliest date for adoption of IFRS to 2015 (SEC 2010).

³ For example, several organizations including Computer Sciences Corporation, PricewaterhouseCoopers, and Financial Executive International made comments indicating that a principles-based approach would result in transactions being accounted for according to their economic substance rather than their form (see the FASB website for these and other comment letters: http://www.fasb.org/jsp/FASB/CommentLetter_C/CommentLetterPage?project_id=1125-001).

whether at least one member of the audit committee is a financial expert. Archival studies have shown audit committee strength to have an effect on the financial reporting process in rules-based standards environments (Klein 2002; Bédard et al. 2004). However, evidence on its effect in more principles-oriented environments and on the interactive effect of standard precision and audit committee strength is lacking (Libby et al. 2002).

To explore these issues, we conduct two experiments in which experienced U.S.-based financial statement preparers are placed in a lease classification decision context. We manipulate standard precision between participants, where more precise and less precise standards are based on FASB Statement No. 13 and International Accounting Standard (IAS) 17, respectively. Audit committee strength is manipulated between participants as either strong or weak.

The results of experiment 1 indicate that financial statement preparers are less likely to report aggressively (i.e., they are more likely to capitalize the lease) when applying a less precise (more principles-based) lease classification criterion than when applying a more precise (more rules-based) criterion. This result provides support for the SEC's and FASB's hope that a move toward more principles-based standards could result in better, or less aggressive, financial reporting. Consistent with prior archival research, we also find that preparers asked to apply a more precise standard are less likely to report aggressively in the presence of a strong audit committee than in the presence of a weak audit committee. However, we find that the influence of a strong audit committee diminishes in the principles-based setting. Interestingly, we find significantly less variability among preparers' reporting decisions when a less precise standard is in place. This suggests that, contrary to the concern of some interested parties, the application of more principles-based standards need not result in less comparability than more precise standards. Experiment 2 confirms the findings of experiment 1 and finds support for a three-path mediation model. Specifically, we find that the less precise the standard, the more concerned preparers are about second-guessing and possible costs imposed through regulation and litigation. This, in turn, results in an increased desire to reflect the underlying economics of transactions and events in the financial statements and, ultimately, in less aggressive reporting.

The findings of our study suggest a number of important implications. For example, these findings should be of interest to policy makers as they contemplate whether and when to adopt IFRS. Our results have encouraging implications for U.S. financial reporting for a move to a more principles-based accounting standards model. Also, our study provides insight regarding recent regulatory changes that address the role and composition of audit committees. Our results suggest that, under certain circumstances, principles-based standards can ease the burden on an audit committee (whether strong or weak) to curtail management's aggressive reporting choices.

Section II provides background information and develops our hypotheses. Sections III and IV describe the experiments used to test hypotheses and present the results. Section V provides a summary and offers conclusions, implications, and suggestions for future research.

II. BACKGROUND AND HYPOTHESES

Principles-Based versus Rules-Based Accounting Standards

The Current Environment

Much of the recent interest in IFRS is predicated on the notion that these standards are principles-based, whereas U.S. GAAP is described as (and often criticized for) being rules-based. Rules-based standards typically provide very detailed guidance with bright-line tests. A perceived benefit of more detailed implementation guidance is greater comparability of financial statements across companies (Schipper 2003). On the other hand, it has been suggested that rules-based standards lead to a "show me where it says I can't" attitude, which, in turn, can lead to dysfunc-

tional financial reporting behavior (Weil 2002, 3). Excessively detailed reporting guidance can invite transaction structuring and incentive-consistent standard interpretation to achieve preferred accounting treatments (FASB 2002; Bockus et al. 2003; Nelson 2003).

The FASB's Statement of Financial Accounting Standards (SFAS) No. 13, *Accounting for Leases*, with its detailed list of criteria for lease classification containing several bright-line thresholds, is cited as the "poster child" for rules-based standards (Maines 2007, 360). In developing SFAS No. 13, the FASB had hoped that explicit lease classification rules would eliminate individual judgment, resulting in consistent application of the standard across firms (Shortridge and Myring 2004). However, because bright-line tests were established, companies have been able to structure and interpret lease contracts to avoid capitalization, which tends to present a more favorable picture of a company's overall financial condition (Imhoff and Thomas 1988; Pulliam 1988). As a result, "the explicit rule allows the off balance sheet financing to continue, *and provides justification for the treatment*" (Shortridge and Myring 2004, 3; emphasis added).

Principles-based standards, which provide limited interpretive and implementation guidance, are the perceived solution to problems caused by rules-based standards. Less guidance, in theory, increases the need to apply professional judgment consistent with the intent of the standards. Both the FASB (2002) and the SEC (2003) believe that this will ultimately result in more meaningful and informative financial statements.⁴

Related Research

Several experimental studies have explored auditor judgments in light of specific accounting standards or guidance (e.g., Gibbins et al. 2001; Ng and Tan 2003; Kadous et al. 2003).⁵ For example, Ng and Tan (2003) find that, in the presence of a weak audit committee, auditors are more likely to allow aggressive revenue recognition when no guidance/rule governs the transaction than when there is authoritative guidance specifying the appropriate reporting treatment. Trompeter (1994) finds that, when authoritative guidance limits the range of acceptable accounting treatments, audit partners are less influenced by client preferences. In an imprecise standard context, Hackenbrack and Nelson (1996) show that auditors approve reporting decisions that are consistent with their incentives and use the vagueness of the imprecise standard to justify their decisions. While these studies provide some evidence as to how *auditors* react to the existence or specificity of authoritative guidance, they assume, either implicitly or explicitly, that managers will always choose to report aggressively. This line of research does not speak to how standard precision influences the level of aggressiveness in reporting decisions of financial statement *preparers*.

Three studies have employed an experimental approach to examine the influence of standard precision on reporting decisions made by financial statement preparers (Hoffman and Patton 2002; Psaros and Trotman 2004; Jamal and Tan 2010). Hoffman and Patton (2002) hypothesize that less precise standards for governmental financial reporting provide more latitude for federal financial

⁴ Specifically, the FASB (2002) concludes that adopting a principles-based approach will result in greater judgment, leading to: accounting treatments that conform to the substance of a transaction, improved transparency, enhanced comparability, increased responsiveness to emerging accounting issues, and facilitation of international financial reporting standard convergence.

⁵ Cuccia et al. (1995) explore rule precision in a *tax* setting, finding that moving from a vague rule to a more precise rule has no effect on tax practitioners' propensity to take an aggressive position on an individual's return. When a vague standard exists, tax preparers use the vagueness in the standard to justify their aggressive tax position; when a precise standard exists, they use the vagueness in facts to justify their position. However, as the authors point out, results from their study of tax practitioners might not generalize to financial reporting decisions because of differences in penalties and professional responsibilities (Cuccia et al. 1995), as well as in the scope of impact (e.g., aggressive reporting in an individual tax setting affects a few people at most, limiting the financial/social impact, while aggressive financial reporting by publicly traded companies can have a much broader and deeper impact).

officers to reach reporting decisions that are aligned with their incentives. Contrary to expectations, they find that replacing a less precise recognition threshold (“probable”) with a more precise threshold (“more likely than not”) does not affect the likelihood that these preparers will recognize a contingent loss. They speculate that the lack of an effect could be due to a weak manipulation, as participants did not appear to perceive the new threshold to be more precise than the original.

Using two separate experiments involving a consolidation judgment, Psaros and Trotman (2004) consider whether corporate accountants will justify aggressive judgments by more aggressively interpreting case-specific information. Results of the first experiment, conducted in a “substance-over-form” (less precise) standard setting, suggest that accountants who make more aggressive judgments tend to interpret case-specific information more aggressively. They obtain similar results in the second experiment, which is set in a more precise standard environment. As the researchers did not include both levels of standard precision in one experiment, they are unable to provide direct evidence on the relative impact that the level of standard precision has on reporting behavior. However, by comparing results from the two experiments, they suggest that aggressive reporting is more likely in rules-based settings. Psaros and Trotman (2004) call for further research designed to more directly compare the relative influence of standards of differing precision on preparers’ reporting decisions.

Jamal and Tan (2010) examine whether auditor type (e.g., principles-oriented or rules-oriented) affects financial managers’ reporting decisions under rules-based and principles-based standards. They find that auditor type has no effect on reporting decisions under a rules-based standard. However, under a principles-based standard, financial managers are less likely to report aggressively when the auditor is principles-oriented. They speculate that improved financial reporting will result only if a move toward more principles-based standards is accompanied by a shift in auditors’ mindsets toward being more principles-oriented. This shift would appear likely to happen relatively quickly because auditors have a responsibility to assure that their clients are in compliance with prevailing standards, regardless of standard type. That is, if there is a move toward principle-based standards such as IFRS, a “rules-based” auditor would quickly have to adjust to an environment in which there are no bright-line thresholds on which to fall back. Thus, the cells from their study that are most relevant for purposes of our study are the two “matched” cells (i.e., principles-oriented auditor/principles-based standard and rules-oriented auditor/rules-based standard). Results from these two cells are not inconsistent with the notion that principles-based standards will typically result in less aggressive financial reporting by preparers.

Standard Precision and Aggressive Financial Reporting

Rules-based standards have developed over time, in part, because of demand for them by financial statement preparers and auditors, who believe that such detailed rules shield them from potential criticism for aggressive reporting (Benston et al. 2006). Aggressive reporting is likely to be more difficult to justify with a less precise standard in which a threshold is not explicitly stated (Maines 2007). The costs associated with being perceived to be out of compliance are similar under both more precise and less precise standards. However, the uncertainty surrounding the risk of being perceived to be out of compliance (i.e., second-guessed) is inherently greater without detailed guidance (Nelson et al. 2002).⁶ Thus, preparers applying less precise standards could decide not to select their desired accounting treatment because of the increased risk of second-guessing and the attendant costs. Theoretical research in the areas of law and economics supports the idea that less precise rules could result in less aggressive decisions (Calfee and Craswell 1984; Craswell and Calfee 1986). This outcome could be more likely in the post-SOX era because of

⁶ This assumes a strong regulatory enforcement environment (e.g., the post-SOX regulatory environment).

required CEO and CFO certification of the financial statements. Following this reasoning, along with a belief that “objectives-oriented standards allow accounting professionals to operationalize accounting treatments in a manner that best fulfills the objective of each standard and thereby best captures the underlying economic reality,” the SEC (2003, 11) and the FASB (2002) have indicated a preference for less precise standards.

There is an implicit assumption underlying this belief that, when the constraints imposed by rules-based standards are removed, financial statement preparers will desire to make accounting choices that best reflect economic reality. However, the FASB (2002, 9) acknowledges that a principles-based approach “could lead to abuse, whereby the principles in accounting standards are not applied in good faith consistent with the intent and spirit of the standards.” There are a number of environmental factors that could influence whether a principles-based standard regime leads to more meaningful financial reporting or to abuses of the less-precise guidance. For example, the level of regulatory scrutiny of accounting treatment choices, the degree of consensus among regulators about the appropriate treatment, and the ability of rules-based standards to constrain aggressive choices could influence the efficacy of a principles-based regime.

We choose to examine a setting that best represents the concerns/complaints levied by investors and regulators. Specifically, we examine a rules-based standard that has been unable to constrain aggressive behavior, even in a strong post-SOX regulatory environment in which there is concern about aggressive reporting choices.⁷ In such a setting, we expect that the level of standard precision will affect the extent to which financial statement preparers are concerned about second-guessing and possible costs imposed through regulation and litigation. Financial statement preparers should be more concerned about second-guessing when standards are less precise (Nelson et al. 2002). This concern will result in an increased desire to reflect the underlying economics of transactions and events in the financial statements. Presumably, if preparers can show that their accounting choices fairly present economic reality, then they will be better able to defend themselves when second-guessed by external parties (e.g., the SEC). Thus, we argue that less precise standards result in an increased concern about second-guessing and related costs, which in turn results in a greater interest in fair presentation. This reasoning leads to the following hypothesis:

- H1:** Financial statement preparers applying a less precise financial reporting standard will be less likely to make an aggressive financial reporting decision than preparers applying a more precise standard.

The Influence of the Audit Committee on Aggressive Reporting in Principles-Based and Rules-Based Settings

The audit committee represents a key corporate governance mechanism, in that it is viewed as one of the most important monitors of the financial reporting process (Blue Ribbon Committee 1999). However, a spate of high-profile corporate accounting scandals (e.g., Enron, WorldCom, Tyco) brought the effectiveness of audit committees into question. Regulators and investors have called for stronger audit committees to help ensure higher quality financial reporting. Congress responded by enacting SOX, which, among other things, expands the responsibilities of the audit committee and places a greater focus on its role in enhancing the financial reporting process. Under SOX, the audit committee now is responsible for hiring the company's independent auditor, overseeing the work of the auditor, resolving financial reporting differences between management and the auditor, and monitoring internal controls (Lander 2004). Concurrent with the expansion of

⁷ This also represents a setting in which we would be likely to find a benefit of a principles-based approach, should such a benefit exist. For example, if a shift to principles-based standards is accompanied by a reduction in the level of regulatory scrutiny, the potential abuses the FASB (2002) acknowledges are more likely to occur.

the audit committee's role, there has been a push toward greater independence and expertise of audit committee members. SOX now requires all members of the audit committee to be independent (i.e., not currently part of management). Companies must now also disclose whether at least one member of the audit committee is a financial expert. Further, while there are no strict guidelines regarding the frequency of audit committee meetings (the Blue Ribbon Committee [1999] recommends at least four meetings per year), some have called for audit committees to be more active and diligent in carrying out their oversight duties (DeZoort et al. 2002; Stewart and Munro 2007).

Thus, conventional wisdom is that audit committee independence, expertise, and activity can enhance the monitoring and oversight of management, leading to less aggressive financial reporting choices. Prior research appears consistent with this notion (Bédard et al. 2004). Results of archival studies examining U.S.-listed firms suggest that independent members are less likely to be sanctioned for fraudulent or misleading financial reporting, and are negatively associated with abnormal accruals, aggressive earnings management, and earnings restatements (Abbott et al. 2000; Klein 2002; Abbott et al. 2004; Bédard et al. 2004). Similarly, financial expertise on the audit committee is negatively associated with discretionary accruals and aggressive earnings management and is viewed favorably by capital markets (Xie et al. 2003; Abbott et al. 2004; Bédard et al. 2004; DeFond et al. 2005). Further, the frequency of audit committee meetings is negatively related to financial reporting misstatements, earnings management, and perceived audit risk (Abbott et al. 2000; Stewart and Munro 2007). Thus, in the rules-based U.S. reporting environment, a strong audit committee can serve as a monitoring mechanism that helps mitigate aggressive financial reporting behavior.

However, it is less clear whether the same relationship exists, or is as strong, in environments that employ principles-based accounting standards. Assuming a relatively strong regulatory environment, preparers using principles-based standards are likely to perceive greater risk of regulator sanctions due to the inherent uncertainty of the standard. This uncertainty could lead to an increased desire to reflect the economic substance of a transaction, in turn lessening the audit committee's burden to constrain aggressive reporting choices. Many countries using the less precise IFRS either do not require audit committees (e.g., Russia) or they are fairly recent developments (e.g., the European Union) and, as a consequence, are likely weak relative to U.S. audit committees. Thus, there is no direct evidence regarding the relationship between audit committee strength and aggressive reporting in a principles-based setting, and what indirect evidence there is does not lead to clear conclusions regarding the existence or strength of the relationship (Collier and Gregory 1996; Peasnell et al. 2000; Chen and Cheng 2007). However, there are certain settings under which inferences could be drawn regarding the influence of audit committee strength in principles-based standard environments.

Recall that, in our setting,⁸ we expect that preparers applying a more principles-based standard will be more concerned about regulator second-guessing and capturing the economic substance of the transaction than preparers applying our rules-based standard, irrespective of the strength of the audit committee (i.e., the H1 expectation). If this is the case, then preparers applying the less precise standard will be more likely than preparers applying the more precise standard to choose the treatment preferred by the audit committee, thereby lessening the burden on the audit committee to curb aggressive reporting. That is, the less precise standard should dampen pursuit of aggressive financial reporting regardless of audit committee strength, reducing the need for the audit committee to exert its influence regarding this reporting choice. We therefore test the following hypothesis:

⁸ Our setting is one in which the rules-based standard has been relatively ineffective at curtailing aggressive reporting, even in a strong post-SOX regulatory environment where there is concern about aggressive reporting choices.

- H2:** The difference in financial reporting judgments made in the presence of a strong audit committee versus a weak audit committee will be greater when a financial reporting standard is more precise than when it is less precise.

III. EXPERIMENT 1

Participants and Design

Participants in experiment 1 are 96 experienced financial statement preparers (55 CFOs, 23 Controllers, and 18 VPs of Finance) with an average of 25.5 years of professional work experience.⁹ As our experiment asks participants to assume the role of a company controller faced with a lease classification decision, it was important that we select experienced executives who are likely to play a key role in the financial reporting decisions of their companies.

To test our hypotheses, we conduct an experiment requiring participants to make a lease classification decision in which the two variables of interest (standard precision and audit committee strength) are manipulated at two levels. The *more precise* condition for standard precision includes a bright-line criterion for lease capitalization (lease term is "equal to 75% or more" of expected life) consistent with SFAS No. 13 (FASB 1976). The *less precise* condition includes a vaguer criterion (lease term is "for the major part" of expected life) based on IAS 17 (IASC 1997). In the *strong audit committee* condition, participants are told that all audit committee members are independent with no disclosed relationship with the company and all qualify as financial experts as defined by the SEC, and that the audit committee meets frequently (11 to 12 times per year). For the *weak audit committee* condition, while all committee members qualify as independent, participants are informed that one member has no disclosed prior relationship with the company (two are former officers of the company), one qualifies as a financial expert, and the committee meets infrequently (two to three times per year).¹⁰

Procedures

We provided participants with a research instrument containing four sections. Section I included guidelines for classifying a lease as either a capital or operating lease. Participants were told only one criterion (relating to the ratio of lease term to estimated economic life) was relevant to the lease classification decision. Specifically, one group of participants was instructed that a lease must be classified as a capital lease if the lease term is "equal to 75% or more" of the

⁹ We mailed instruments to 1,000 individuals identified by the American Institute of Certified Public Accountants (AICPA) as financial executives. We received replies from 106 individuals and 63 were returned as undeliverable. The resulting response rate of 11.31 percent (106 responses divided by 937 delivered) is consistent with prior studies involving CFO/controller participants (e.g., Graham and Harvey 2001; Gibbins et al. 2007; Sanchez et al. 2007). There were ten unusable responses: seven instruments were completed by inappropriately classified individuals such as staff accountants, payroll clerks, and tax accountants (our conclusions remain the same with or without these individuals), and three were returned with no response on the dependent variable. Thus, 96 usable responses remained. Comparisons of early and late responders indicate no significant difference, suggesting that nonresponse bias does not drive our results. With respect to the demographics of our participants (e.g., current or prior experience working at a publicly traded company or as an external auditor, experience with leases), there were no significant differences between conditions and, when included in our analyses, the demographic variables were neither significant nor altered the conclusions we draw.

¹⁰ The features present in both our strong and weak audit committee conditions were selected to represent audit committees that can, and do, exist in today's regulatory environment. For example, our weak committee meets the SOX requirement of independence (no members are currently affiliated with the company) and recommendation for expertise (one member qualifies as a financial expert). With respect to independence, there are numerous publicly traded companies with former employees serving on their audit committees (e.g., FedEx, Goldman Sachs, Kohl's, and Sunoco). Similarly, many firms have a sole designated financial expert (Carcello et al. 2006), including H.J. Heinz Co., CSX Corp., Texas Instruments, Inc., and Nike, Inc. Further, as there are no requirements for frequency of meetings, there is much between-firm variation. Consistent with our weak audit committee manipulation, a number of firms' audit committees meet infrequently (i.e., four or fewer meetings; see Huron 2006).

estimated economic life of the leased property (the more precise standard based on SFAS No. 13). Another group was instructed that a lease must be capitalized if the lease term is “for the major part” of the estimated economic life of the leased property (the less precise standard based on IAS 17). Both groups were told that a lease must be classified as an operating lease if it does not meet the capital lease criterion. In addition, participants were provided the following definitions of “lease term” and “bargain renewal option” based on definitions in both SFAS No. 13 and IAS 17:

- *Lease term* is defined as the fixed non-cancelable term of the lease plus all periods covered by bargain renewal options.
- *Bargain renewal options* allow the lessee to renew the lease for a rental sufficiently lower than the fair rental of the property such that exercise of the option appears, at the inception of the lease, to be reasonably assured.

Given the definition of lease term, lease classification requires financial statement preparers to first judge whether any renewal option embedded in the lease is a bargain. Second, preparers must judge whether the lease term, i.e., the number of years in the fixed non-cancelable portion of the lease plus the number of years in any bargain renewal option, meets the threshold established in the standard.

Section II asked participants to assume the role of controller of a hypothetical company that has entered into a lease for new equipment with an estimated economic life of ten years. The lease has a non-cancelable lease period of seven years. At the end of the initial non-cancelable lease period, the agreement provides the company the option to renew the lease for an additional year, with the monthly rental payment set at a rate that allows for some discretion in judging whether the renewal option represents a bargain.¹¹ Participants applying the “equal to 75% or more” criterion must exercise their judgment in measuring the “lease term” by determining whether the lease renewal option represents a bargain. They can justify an operating lease classification by arguing that the renewal option does not represent a bargain and therefore the lease term (seven years) is only 70 percent of the asset’s economic life (ten years). However, if they judge the renewal option to be a bargain, then they should add the bargain renewal option period (one year) to the initial lease period (seven years) and view the lease term as eight years, which is 80 percent of the asset’s economic useful life. In this case, the bright-line threshold in the standard should cause them to classify the lease as a capital lease. Participants applying the less precise standard must exercise similar judgment in determining whether the lease renewal option represents a bargain. However, their interpretation of “for the major part” also factors into their lease classification decisions.¹²

Participants were provided a summary of the differential impact of the two accounting treatments on the company’s financial statements and key ratios at the end of the first year of the lease. The summary shows that capitalization of the lease generally produces less favorable financial results, including lower EPS and return on assets figures and a larger debt-to-equity ratio. Thus, participants generally should have an incentive to classify the lease as an operating lease. To make this incentive more salient, participants were told that the company is publicly traded and that it

¹¹ Controllers at several *Fortune* 500 companies suggested that setting the renewal option at approximately 90 percent of the fair rental value would be sufficiently ambiguous to require judgment as to whether it represents a bargain (i.e., whether to include it as part of the total lease term). Thus, we set the option at 90 percent to provide an opportunity for the financial incentives of alternative classification to factor into participants’ lease classification decisions.

¹² That is, for the renewal option to place a participant applying the less precise standard in a similar professional decision context as those applying the more precise standard, a less precise standard participant must interpret “for the major part” as ≥ 71 and ≤ 80 percent of economic life. We, therefore, conduct an analysis comparing those participants in the 71 to 80 percent range to those in the more precise group (i.e., for both groups of participants, the judgment regarding the option results in a different classification). We discuss this in greater detail below.

intends to raise capital in the second year of the lease term through the issuance of both common stock and the sale of bonds. After considering the case information, participants recorded their lease classification decisions and responded to several post-experimental and demographic questions (Sections III and IV of the instrument).

Experiment 1 Results

Our hypotheses are tested using a 2×2 ANOVA (standard precision by audit committee strength) with the financial statement preparer's lease classification decision serving as the dependent variable.¹³ Due to the directional nature of expectations, all tests of hypotheses are one-tailed. Cell means and ANOVA results are presented in Table 1.¹⁴

Standard Precision

H1 predicts that financial statement preparers applying a precise standard will be more likely to make an aggressive financial reporting decision (i.e., classify the lease as an operating lease) than preparers applying an imprecise standard. Preparers recorded their lease classification decisions on a ten-point, forced-choice scale (1 = "Definitely classify as an operating lease" and 10 = "Definitely classify as a capital lease"). Table 1 reports results consistent with expectations. Specifically, preparers are more likely to report aggressively when presented with a more precise standard (mean = 4.98) than with a less precise standard (mean = 7.83, $F = 23.36$, $p < 0.001$).

It is important to establish that these findings are not solely a function of our manipulation of precision resulting in essentially different decision contexts for the less and more precise conditions. That is, for participants in the less precise standard condition, the term "for the major part" is subject to interpretation. For a more meaningful comparison, participants in the less precise condition must be placed in a similar decision context to those in the more precise condition. Specifically, participants' interpretations of "for the major part" must place them in a situation in which their professional judgment regarding the renewal option is meaningful (i.e., the judgment regarding the option results in a different lease classification decision).¹⁵ Of the 47 participants in the less precise group, 19 indicated that they interpreted the phrase within this professional judgment range (i.e., as ≥ 71 percent and ≤ 80 percent).¹⁶ As a more restrictive test of the effect of

¹³ As a manipulation check for audit committee strength, participants were asked to assess (on a seven-point scale where 1 = Low and 7 = High) the company's audit committee along three dimensions: independence from management, accounting/financial expertise, and frequency of meetings. t-tests indicate that participants in the strong audit committee condition assessed the committee's independence (mean = 6.29), accounting/financial expertise (mean = 6.41), and frequency of meetings (mean = 6.63) to be significantly higher (all p 's < 0.001) than participants in the weak condition (means of 2.93, 2.52, and 2.84 for independence, expertise, and frequency of meetings, respectively), suggesting that our audit committee manipulation was successful. With respect to the standard precision manipulation, 85 percent of participants correctly identified the applicable standard. Our conclusions remain the same if participants who failed this manipulation check are excluded from the analysis.

¹⁴ Levene's test for equality of variances indicates violations of the ANOVA assumption of homogeneity of variances for the lease decision (discussed later). While ANOVA results are typically quite robust to violations of this assumption (Box 1954; Lindman 1974), we also conduct analyses on transformations (i.e., square root, natural log, inverse function, and ranking transformations) of our dependent variable for assurance. Results of these analyses are consistent with the results presented in Table 1.

¹⁵ That is, if a participant in the less precise condition interprets "for the major part" as ≤ 70 percent, then the fixed term of the lease alone would automatically result in capitalization. If the interpretation is > 80 percent, then the lease would be treated as an operating lease regardless of whether the renewal option is determined to be a bargain. In either case, the participant's professional judgment regarding the renewal option would be irrelevant to the classification of the lease. Thus, we focus this analysis on those participants whose responses were within the range that requires a "professional judgment" similar to that required of participants in the more precise standard condition.

¹⁶ Mean and median responses for the 47 participants in the less precise group are 69.13 and 70 percent, respectively, with 27 participants interpreting "major part" as ≤ 70 percent, 19 as within the 71 to 80 percent "judgment required" range, and one as > 80 percent.

TABLE 1

Experiment 1: Lease Classification Decisions (Likelihood Scale)

Panel A: Mean (Standard Deviation)^a

Standard Precision	Audit Committee		Overall
	Weak	Strong	
Less precise	8.14 (2.24) n = 21	7.58 (2.75) n = 26	7.83 (2.52) n = 47
More precise	4.13 (3.08) n = 23	5.73 (3.51) n = 26	4.98 (3.38) n = 49
Overall	6.05 (3.36) n = 44	6.65 (3.25) n = 52	

Panel B: ANOVA Results

	F	p-value ^b
Model	8.76	<0.001
Experimental Variables		
Standard Precision	23.36	<0.001
Audit Committee	0.73	0.396
Interaction		
Standard Precision × Audit Committee	3.19	0.039

^a Participants indicated the likelihood they would classify a lease as either an operating lease or a capital lease using a ten-point scale numbered from 1 ("Definitely classify as an operating lease") to 10 ("Definitely classify as a capital lease").

^b Where expectations are directional, p-values are based on one-tailed tests.

standard precision, we retest H1 using only these 19 participants (non-tabulated mean = 7.00 for these individuals). A comparison with responses from the more precise group (mean = 4.98) provides additional support for H1 ($p = 0.013$).¹⁷

To further explore the effect of standard precision, we dichotomize preparers' responses at the midpoint as a measure of the practical significance of their lease classification decisions. Analysis with the dichotomized variable reveals that participants in the more precise standard condition are significantly more likely to take the aggressive position of classifying the lease as an operating lease than those applying a less precise standard. Sixty-one percent of preparers in the more precise condition classified it as an operating lease versus 15 percent in the less precise condition ($\chi^2 = 21.74$, $p < 0.001$; Table 2), providing further support for H1.

Additionally, in accordance with the development of H1, participants appear to interpret the facts of the lease scenario in a manner consistent with their classification decisions. We find a significant correlation between their classification decisions and their interpretation of whether the

¹⁷ Results are similar when the data are parsed to include only those in the less precise condition who interpret "for the major part" as ≥ 75 percent (i.e., participants whose interpretations are as high as, or higher than, the threshold in the more precise standard).

TABLE 2
Experiment 1
Contingency Analysis of Lease Classification Decision by Standard Precision Condition

Standard Precision	Lease Classification Decision		Row Total
	Operating Lease	Capital Lease	
Less precise	7	40	47
Row %	14.89%	85.11%	100.00%
More precise	30	19	49
Row %	61.22%	38.78%	100.00%
Column total	37	59	

$$\chi^2 = 21.74 \text{ (} p < 0.001 \text{)}$$

Participants' ten-point likelihood responses were dichotomized at the midpoint such that responses of 1 to 5 on the scale are classified as *Operating Lease* and responses of 6 to 10 are classified as *Capital Lease*.

option to renew the lease at the end of the lease term represents a bargain (Pearson $r = 0.580$; $p < 0.001$). That is, participants who indicate they are more likely to capitalize the lease tend to believe more strongly that the renewal option represents a bargain.

The Influence of the Audit Committee

H2 predicts that audit committee strength will have a greater effect on aggressive reporting in a rules-based regime (where a strong audit committee could inhibit financial statement preparers' opportunistic application of a standard) than in a more principles-based regime (where greater concern about regulator sanctions could lessen the burden on the audit committee to constrain aggressive reporting). Table 1 reports a significant and directionally consistent interaction between standard precision and audit committee strength ($F = 3.19$, $p = 0.039$), providing support for H2. Further, non-tabulated comparisons within each of the standard precision conditions reveal that audit committee strength influences lease classification decisions under the rules-based regime (means = 5.73 and 4.13 for the strong and weak audit committee conditions, respectively, $p = 0.049$), but not under the principles-based regime (means = 7.58 and 8.14 for the strong and weak audit committee conditions, respectively, $p = 0.441$).

An interesting finding emerges when we compare the mean responses in Table 1 in the more precise standard/strong audit committee condition (mean = 5.73) with those in the two less precise standard conditions (means = 7.58 and 8.14 for the less precise/strong and less precise/weak conditions, with two-tailed p -values of 0.040 and 0.007, respectively). We find that both of the less precise standard conditions lead to less aggressive reporting, regardless of audit committee strength. While it is important to stress the context-specific nature of this result, the shift to a more principles-based standard appears to have a greater dampening effect on aggressive reporting than increasing the strength of the audit committee does in our setting.

Comparability

A concern with moving from more precise (rules-based) to less precise (principles-based) standards is that it could result in reduced financial statement comparability across firms (FASB 2002; Nelson 2003; Schipper 2003; SEC 2003). If this concern has merit, one would expect the variability of lease classification decisions to be greater with a less precise standard than with a more precise standard. Contrary to concerns, Table 1, Panel A reveals that there is *less* variability

(as measured by the standard deviation) in the decisions made by preparers in the less precise standard group than in the more precise group ($SD = 2.52$ and 3.38 , respectively). Levene's test for equality of variances indicates a significant difference between the two groups ($F = 15.397$, $p < 0.001$, two-tailed, non-tabulated). This suggests that application of more principles-based standards need not result in less inter-firm comparability than more precise standards.

IV. EXPERIMENT 2

Concern for Sanctions and Economic Substance: A Three-Path Mediating Model

In experiment 1, we find a significant relationship between standard precision and financial statement preparers' lease classification decisions. Specifically, a less precise (more principles-based) standard results in less aggressive financial reporting (H1). The development of H1 suggests two potential mediators of this relationship. With their bright-line thresholds, rules-based standards can invite transaction structuring and incentive-consistent standard interpretation to achieve preferred accounting treatments (FASB 2002; Bockus et al. 2003; Nelson 2003). Thus, as Maines (2007) suggests, financial statement preparers could prefer rules-based standards because they are believed to reduce second-guessing by regulators as to whether a standard has been applied opportunistically. Aggressive reporting could be more difficult with a principles-based standard in which a threshold is not explicitly stated, as the risk of being perceived to be out of compliance is greater due to the inherent uncertainty of the standard (Nelson et al. 2002). When there is uncertainty regarding the proper accounting, preparers might decide not to select their desired treatment because of the increased risk of second-guessing by regulators, as well as the costs accompanying such sanctions (Calfee and Craswell 1984; Craswell and Calfee 1986). This increased concern for second-guessing would result in a greater desire to reflect the underlying economics of transactions and events in the financial statements. That is, if preparers can show that their reporting decisions fairly represent economic reality, they will be better able to defend themselves when second-guessed by regulators. We propose a three-path (two-mediator) model in which concern about second-guessing and interest in reporting the economic substance of a transaction sequentially mediate the relationship between standard precision and aggressive reporting. Experiment 2 is designed to explore this mediating relationship and tests the following hypothesis:

H3: Concern about second-guessing by regulators and the desire to report the economic substance of the transaction will (sequentially) mediate the relationship between standard precision and financial statement preparers' aggressive financial reporting decisions.

Task and Procedure

Similar to experiment 1, participants were 92 experienced financial statement preparers (i.e., CFOs, Controllers, VPs of Finance) with an average of 24.9 years of professional work experience.¹⁸ Experimental materials were similar to those used in experiment 1. Again, we manipulate standard precision and audit committee strength and ask participants to make a lease classification decision. However, we made minor modifications to the details of the lease agreement in an effort to rule out the possibility that a ceiling effect is responsible for our experiment 1 finding that the burden on the audit committee to curb aggressive reporting eases with a move toward a more principles-based standard (i.e., mean responses for participants in both principles-

¹⁸ We conducted another mailing to 1,000 financial executives, receiving 99 replies with 88 returned as undeliverable for a response rate of 10.9 percent. Seven responses were unusable: five were completed by inappropriately classified individuals (our conclusions remain the same with or without these individuals), and two were returned with no response on the dependent variable, resulting in 92 usable responses. Demographic variables (e.g., experience-related measures, years at current job) were not significantly different between conditions and, when included in our analyses, were neither significant nor altered the conclusions we draw.

based conditions were relatively close to the high, or “capitalize,” endpoint of the response scale). The modifications to the lease agreement were designed to allow for greater freedom for upward movement on the scale, such that an effect of audit committee strength would be more easily detectable. Specifically, relative to experiment 1, we reduced the fixed non-cancellable term of the lease to five years. The option to renew the lease was increased to three years, but at a rental payment of 92 percent of the equipment’s fair rental value. The rental payment was increased to make it easier for participants to conclude that the renewal option is *not* a bargain and thus justify classifying the lease as an operating lease (i.e., to move participants down the response scale to, hopefully, rule out the possibility of a ceiling effect driving our H1 findings for experiment 1). Further, the changes in the length of the fixed term and option increase the proportion of participants in the principles-based standard condition for whom the lease classification requires judgment. These modifications allow for a wider range of participants’ interpretations of “for the major part” (i.e., ≥ 51 percent and ≤ 80 percent) to place them in a situation in which their professional judgment regarding the renewal option is meaningful (i.e., the judgment regarding the option results in a different lease classification decision).

In an effort to attempt to identify the mechanisms driving our findings in experiment 1 (and test H3), we ask about several factors that could influence participants’ lease classification decisions. Participants were asked, post-experimentally in experiment 2, to consider how much their decisions were influenced by five potentially influential factors (relative to other potential factors). These factors relate to their desire to: (1) “Report the economic substance of the lease in the financial statements,” (2) “Avoid possible second-guessing of my decision by the company’s audit committee,” (3) “Avoid possible second-guessing of my decision by the company’s auditor,” (4) “Avoid possible second-guessing of my decision by external watchdogs such as the Securities and Exchange Commission,” and (5) “Present the company’s financial position and profitability as favorably as the circumstances will allow.”

Experiment 2 Results

The 2×2 ANOVA results presented in Table 3 confirm the results of experiment 1. Again, we find a significant main effect of standard precision (H1), with preparers applying a more precise standard reporting more aggressively (mean = 3.41) than those applying a less precise standard (mean = 6.19, $p < 0.001$).¹⁹ Dichotomizing participants’ responses, we find that participants in the more precise standard condition are significantly more likely to take the aggressive position of classifying the lease as an operating lease (77.3 percent) than those applying a less precise standard (33.3 percent; $\chi^2 = 17.86$, $p < 0.001$, non-tabulated), consistent with H1. Further, we again report significantly less variability in decisions made by preparers in the less precise standard group than in the more precise group ($SD = 1.48$ and 2.20 , respectively; Levene Statistic = 14.918 , $p < 0.001$). This result provides additional evidence that principles-based standards might not reduce inter-firm comparability as some have feared. We again find a significant interaction ($p = 0.012$) between standard precision and audit committee strength (H2). Note also that these results appear to rule out the possibility that a ceiling effect is responsible for our H2 results in experiment 1. That is, although the mean responses in the less precise standard conditions were

¹⁹ All participants in the less precise condition interpreted “for the major part” within the range in which their professional judgment regarding the renewal option is required (i.e., ≥ 51 percent and ≤ 80 percent for experiment 2; overall mean and median responses are 68.92 and 70 percent, respectively). Thus, it is unnecessary to conduct the more restrictive test of H1 we utilized for experiment 1. Further, results are similar when the data are parsed to include only those in the less precise condition who interpret “major part” as ≥ 75 percent (i.e., as high as, or higher than, the threshold in the more precise standard). Eighteen of the 48 participants in the less precise condition fall into this category.

TABLE 3
Experiment 2
Lease Classification Decisions (Likelihood Scale)

Panel A: Mean (Standard Deviation)^a

<u>Standard Precision</u>	<u>Audit Committee</u>		<u>Overall</u>
	<u>Weak</u>	<u>Strong</u>	
Less precise	6.13 (1.46) n = 23	6.24 (1.54) n = 25	6.19 (1.48) n = 48
More precise	2.50 (1.79) n = 22	4.32 (2.23) n = 22	3.41 (2.20) n = 44
Overall	4.36 (2.44) n = 45	5.34 (2.11) n = 47	

Panel B: ANOVA Results

	<u>F</u>	<u>p-value^b</u>
Model	22.76	<0.001
Experimental Variables		
Standard Precision	56.48	<0.001
Audit Committee	6.81	0.011
Interaction		
Standard Precision × Audit Committee	3.13	0.012

^a Participants indicated the likelihood they would classify a lease as either an operating lease or a capital lease using a ten-point scale numbered from 1 ("Definitely classify as an operating lease") to 10 ("Definitely classify as a capital lease").

^b Where expectations are directional, p-values are based on one-tailed tests.

a considerable distance away from the "capital lease" endpoint of the scale (overall mean of 6.19 on a ten-point scale), our finding that the influence of audit committee strength diminishes with a move toward a more principles-based standard still holds.

Table 4 presents means for the regulator second-guessing and economic substance variables used to test H3, as well as three other potential influencing factors. Participants recorded on 11-point scales (0 = "Little influence relative to other factors" and 10 = "Very strong influence relative to other factors") the extent to which their lease classification decisions were influenced by these factors. Consistent with our arguments leading to H3, we find that participants applying a less precise standard are more concerned with avoiding regulator second-guessing when making their lease classification decision than those applying a more precise standard (means = 6.64 and 2.14, respectively, $p < 0.001$). Participants in the less precise standard condition also are more concerned with reporting the economic substance of the lease in the financial statements than participants in the more precise standard condition (means of 7.17 and 4.30, respectively, $p < 0.001$).

H3 proposes a three-path mediation model in which concern about second-guessing by regulators and desire to report the economic substance of a transaction sequentially mediate the rela-

TABLE 4
Experiment 2
Factors Influencing the Lease Classification Decision

Dependent Variable	Less Precise Standard			More Precise Standard		
	Weak Audit Committee (n = 23)	Strong Audit Committee (n = 25)	Total (n = 48)	Weak Audit Committee (n = 22)	Strong Audit Committee (n = 22)	Total (n = 44)
<i>Regulator Second-Guessing</i>						
Mean	6.67	6.60	6.64	1.95	2.32	2.14
(Std. Dev.)	(2.15)	(1.94)	(2.01)	(0.99)	(1.17)	(1.09)
<i>Economic Substance</i>						
Mean	7.09	7.24	7.17	3.77	4.82	4.30
(Std. Dev.)	(2.04)	(1.74)	(1.87)	(1.97)	(1.84)	(1.96)
<i>Audit Committee Second-Guessing</i>						
Mean	1.83	3.44	2.67	1.77	6.00	3.89
(Std. Dev.)	(0.94)	(1.45)	(1.46)	(1.02)	(1.95)	(2.63)
<i>Auditor Second-Guessing</i>						
Mean	3.65	3.60	3.63	3.68	3.91	3.80
(Std. Dev.)	(1.23)	(1.19)	(1.20)	(1.32)	(1.27)	(1.29)
<i>Favorable Financial Presentation</i>						
Mean	3.50	3.28	3.39	7.32	6.41	6.86
(Std. Dev.)	(2.38)	(2.73)	(2.56)	(2.32)	(1.92)	(2.15)

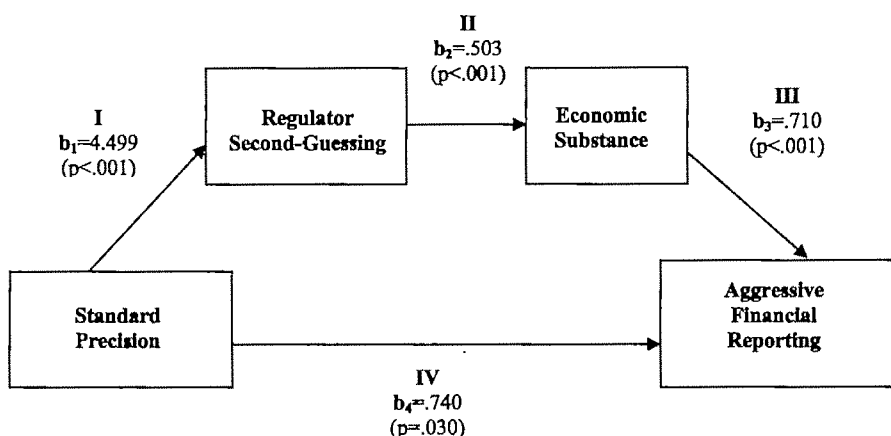
Participants responded to the prompt "Relative to other factors, how much was your lease classification decision influenced by your desire to ..." on 11-point scales, where 0 = "Little influence relative to other factors" and 10 = "Very strong influence relative to other factors" for each of the following factors: "Avoid possible second-guessing of my decision by external watchdogs such as the Securities and Exchange Commission" (*Regulator Second-Guessing*); "Report the economic substance of the lease in the financial statements" (*Economic Substance*); "Avoid possible second-guessing of my decision by the company's audit committee" (*Audit Committee Second-Guessing*); "Avoid possible second-guessing of my decision by the company's auditor" (*Auditor Second-Guessing*); and "Present the company's financial position and profitability as favorably as the circumstances will allow" (*Favorable Financial Position*).

tionship between standard precision and aggressive reporting. Figure 1 reports the results of our test of this model, which support H3. The predicted mediating relationship is demonstrated by the following regression results (MacKinnon and Dwyer 1993):

- (1) Standard precision significantly affects concern for regulator second-guessing (path I in Figure 1, $b_1 = 4.499$, $p < 0.001$),
- (2) Concern for regulator second-guessing influences desire to present the economic substance of the transaction (path II, $b_2 = 0.503$, $p < 0.001$),
- (3) Concern for reporting the economic substance of the transaction significantly influences aggressive financial reporting (path III, $b_3 = 0.710$, $p < 0.001$), and
- (4) Inclusion of concern for reporting the economic substance of the transaction in the analysis reduces the effect of standard precision on aggressive reporting (path IV, $b_4 = 0.740$ and $p = 0.030$ when concern for economic substance is included in the regression versus $b_4 = 2.778$ and $p < 0.001$ when concern for economic substance is *not* included in the regression).

Further, we find that the mediating effect predicted by H3 is significant ($z = 2.35$, $p = 0.009$; see MacKinnon and Dwyer 1993). Thus, this analysis identifies two mechanisms that help explain the

FIGURE 1
Experiment 2: Three-Path Mediation Model and Results



Standard Precision is the treatment variable manipulated at two levels (0 = more precise; 1 = less precise).

Aggressive Financial Reporting is a participant's lease classification decision recorded on a ten-point scale, where 1 = "Definitely classify as an operating lease" and 10 = "Definitely classify as a capital lease." Lower responses are indicative of more aggressive reporting decisions.

Regulator Second-Guessing is a participant's response to the post-experimental question "Relative to other factors, how much was your lease classification decision influenced by your desire to: Avoid possible second-guessing of my decision by external watchdogs such as the Securities and Exchange Commission" (recorded on an 11-point scale where 0 = "Little influence relative to other factors" and 10 = "Very strong influence relative to other factors").

Economic Substance is a participant's response to the post-experimental question "Relative to other factors, how much was your lease classification decision influenced by your desire to: Report the economic substance of the lease in the financial statements" (recorded on an 11-point scale where 0 = "Little influence relative to other factors" and 10 = "Very strong influence relative to other factors").

b_1 through b_4 are the estimated coefficients from the following regression equations:

$$\text{Regulator Second-Guessing} = \beta_{01} + \beta_1(\text{Standard Precision}) + \varepsilon_1$$

$$\text{Economic Substance} = \beta_{02} + \beta_2(\text{Regulator Second-Guessing}) + \varepsilon_2$$

$$\text{Aggressive Financial Reporting} = \beta_{03} + \beta_3(\text{Economic Substance}) + \beta_4(\text{Standard Precision}) + \varepsilon_3$$

To determine the significance of the mediating relationship (paths I through III), a z-score is calculated using the coefficient values and their standard errors. We calculate the variance of the mediating effect ($S_{b_1b_2b_3}$) using a first-order Taylor series estimate (multivariate delta method) (MacKinnon and Dwyer 1993): $z = (b_1b_2b_3) / S_{b_1b_2b_3} = 2.35$ ($p = .009$).

results we obtained in experiment 1 related to the influence of level of standard precision on aggressive financial reporting (H1).

Other Factors Influencing the Lease Classification Decision

Table 4 presents three other potential influencing factors: concern for second-guessing by the audit committee, concern for second-guessing by the auditor, and desire to present the company's

financial position as favorably as the circumstances allow. Consistent with prior archival work in rules-based settings, audit committee strength appears to have significant influence over participants' decisions when applying a more rules-based standard (means of more precise standard/strong audit committee and more precise standard/weak audit committee = 6.00 and 1.77, respectively, $p < 0.001$).

It is interesting to note, however, that a stronger audit committee appears relatively less influential in our more principles-oriented setting (means of less precise standard/strong audit committee condition and more precise standard/strong audit committee condition = 3.44 and 6.00, respectively, $p < 0.001$). This is consistent with our finding that our principles-based standard eases the burden on audit committees (whether strong or weak) to curtail management's aggressive reporting choices.²⁰ That is, when applying our more principles-oriented standard, it appears that preparers are concerned primarily about regulator second-guessing. This, in turn, dampens pursuit of aggressive financial reporting and results in reporting choices that are more consistent with audit committee preferences.

Further examination of potential influential factors in the two more precise standard cells suggests that, when the audit committee is weaker, financial statement preparers are more concerned with presenting the company's financial position as favorably as the circumstances will allow (mean = 7.32) than when the audit committee is stronger (mean = 6.41, $p = 0.082$). It also is interesting to note that presenting a favorable financial position has a much smaller influence on the lease classification decision for participants in the less precise standard condition than for participants in the more precise standard condition (means of 3.39 and 6.86, respectively, $p < 0.001$).

V. SUMMARY AND CONCLUSIONS

As a result of recent corporate accounting scandals, the U.S. Congress passed SOX in an effort to reform the U.S. financial reporting system. In turn, this legislation has pushed the SEC and the FASB to consider reforms designed to inhibit aggressive financial reporting. These reforms include a shift toward more principles-based accounting standards and enhancing the role of the audit committee. We provide the first evidence on how these two SOX-related regulatory initiatives jointly impact financial statement preparers' reporting judgments. Specifically, our study reports the results of two experiments investigating the effect of financial reporting standard precision on experienced financial statement preparers' financial reporting judgments, as well as the potential role that audit committee strength plays in mitigating aggressive financial reporting.

Consistent with expectations, our results show that financial statement preparers are less likely to report aggressively when applying a less precise financial reporting standard than when applying a more precise standard. We also find a significant interactive effect of audit committee strength and standard precision on preparers' lease classification decisions. We find that audit committee strength affects aggressive reporting in a rules-based regime by inhibiting financial statement preparers' opportunistic application of a standard. In contrast, audit committee strength has no effect in a more principles-based regime where preparers are more concerned about regulator second-guessing and reporting the economic substance of a transaction regardless of audit committee strength. We also find support for a three-path mediating model in which concern about second-guessing by regulators and the desire to report the economic substance of a transaction sequentially mediate the relationship between standard precision and preparers' aggressive financial reporting. Further, we find significantly less variability among preparers' financial reporting

²⁰ In contrast to the audit committee (a manipulated variable), we did not expect, nor did we find, significant differences between any of the four cells with respect to concern for second-guessing by the auditor ($p = 0.847$).

decisions when a less precise standard is in place. This suggests that, contrary to the concern of some, the application of more principles-based standards need not result in less inter-firm comparability than more precise standards.

Our findings have implications for both practice and research. With respect to practice, our results provide insight into the principles-versus-rules debate and should be of interest to U.S. policy makers as they continue to contemplate a switch from rules-based to principles-based accounting standards. Our findings suggest that moving toward more principles-based standards (such as the SEC's proposed adoption of IFRS) will not necessarily open the door to greater opportunistic reporting by financial statement preparers; instead, this shift could result in more economically meaningful reporting. Also, our results regarding variability in financial statement preparer responses can help allay concerns regarding inter-firm comparability in a principles-based standards regime. Further, in our setting, we find that switching to a more principles-based approach appears to have a greater dampening effect on aggressive reporting than does strengthening the audit committee. This result, while context-specific, warrants further inquiry by policy makers and researchers. Future research could further explore the relationship between audit committee strength and financial reporting quality under more principles-oriented regimes. Finally, prior research suggests that audit partners believe principles-based standards will reduce their power in resolving auditor-client conflicts (Gibbins et al. 2001). However, our results suggest that the number of such disagreements could diminish, which should help alleviate concerns of auditors' lost leverage.

It is important to note that our experiments focus on a setting in which the rules-based standard has been unable to curb aggressive reporting (i.e., the lease capitalization decision), even given a strong regulatory environment in which there is general concern over aggressive reporting choices. To broaden the generalizability of our findings, future research could explore other settings involving more restrictive rules that more effectively bind financial statement preparers to a conservative treatment. Further research could also explore the effect of differing levels of regulatory scrutiny on principles-based standards. Such research would further our understanding of the roles that standard precision and audit committee strength play in mitigating aggressive financial reporting practices.

REFERENCES

- Abbott, L., J. Park, and S. Parker. 2000. The effects of audit committee activity and independence on corporate fraud. *Managerial Finance* 26 (11): 55–67.
- , S. Parker, and G. F. Peters. 2004. Audit committee characteristics and restatements. *Auditing: A Journal of Practice & Theory* 23 (1): 69–87.
- Blue Ribbon Committee. 1999. *Report and Recommendations of the Blue Ribbon Committee on Improving the Effectiveness of Corporate Audit Committees*. New York, NY: NYSE and NASD.
- Bédard, J., S. M. Chtourou, and L. Courteau. 2004. The effect of audit committee expertise, independence, and activity on aggressive earnings management. *Auditing: A Journal of Practice & Theory* 23 (2): 13–35.
- Benston, G. J., M. Bromwich, and A. Wagenhofer. 2006. Principles-versus rules-based accounting standards: The FASB's standard setting strategy. *Abacus* 42 (2): 165–188.
- Bokus, K. A., W. D. Northcut, and M. E. Zmijewski. 2003. Accounting and disclosure issues in structured finance. In *Corporate Aftershock: The Public Policy Lessons from the Collapse of Enron and Other Major Corporations*, edited by C. L. Culp and W. A. Niskanen, 193–208. Hoboken, NJ: John Wiley.
- Box, G. E. P. 1954. Some theorems on quadratic forms applied in the study of analysis of variance problems, I. Effect of inequality of variance in the one-way classification. *Annals of Mathematical Statistics* 25: 290–302.

- Calfee, J. E., and R. Craswell. 1984. Some effects of uncertainty on compliance with legal standards. *Virginia Law Review* 70 (5): 965–1003.
- Carcello, J. V., C. W. Hollingsworth, and T. L. Neal. 2006. Audit committee financial experts: A closer examination using firm designations. *Accounting Horizons* 20 (4): 351–373.
- Chen, J. J., and P. Cheng. 2007. Corporate governance and the harmonization of Chinese accounting practices with IFRS practices. *Corporate Governance* 15 (2): 284–293.
- Collier, P., and A. Gregory. 1996. Audit committee effectiveness and the audit fee. *European Accounting Review* 5 (2): 177–198.
- Craswell, R., and J. E. Calfee. 1986. Deterrence and uncertain legal standards. *Journal of Law Economics and Organization* 2 (2): 279–303.
- Cuccia, A., K. Hackenbrack, and M. Nelson. 1995. The ability of professional standards to mitigate aggressive reporting. *The Accounting Review* 70 (2): 227–248.
- DeFond, M. L., R. N. Hann, and X. Hu. 2005. Does the market value financial expertise on audit committees of boards of directors? *Journal of Accounting Research* 43 (2): 153–193.
- DeZoort, F. T., D. R. Hermanson, D. S. Archambeault, and S. A. Reed. 2002. Audit committee effectiveness: A synthesis of the empirical audit committee literature. *Journal of Accounting Literature* 21: 38–75.
- Financial Accounting Standards Board (FASB). 1976. *Accounting for Leases*. Statement of Financial Accounting Standards No. 13. Stamford, CT: FASB.
- . 2002. *Proposal—Principles-Based Approach to U.S. Standard Setting*. Norwalk, CT: FASB.
- Gibbins, M., S. Salterio, and A. Webb. 2001. Evidence about auditor-client management negotiation concerning client's financial reporting. *Journal of Accounting Research* 39 (3): 535–563.
- , S. A. McCracken, and S. E. Salterio. 2007. The chief financial officer's perspective on auditor client negotiations. *Contemporary Accounting Research* 24 (2): 387–422.
- Graham, J. R., and C. R. Harvey. 2001. The theory and practice of corporate finance: Evidence from the field. *Journal of Financial Economics* 60: 187–243.
- Hackenbrack, K., and M. W. Nelson. 1996. Auditors' incentives and their application of financial accounting standards. *The Accounting Review* 71 (1): 43–59.
- Hoffman, V. B., and J. M. Patton. 2002. How are loss contingency accruals affected by alternative financial reporting criteria and incentives? *Journal of Accounting and Public Policy* 21: 151–167.
- Huron Consulting Group (Huron). 2006. *2006 Audit Committee Research Report*. Chicago, IL: Huron Consulting Services LLC.
- Imhoff, E., and J. Thomas. 1988. Economic consequences of accounting standards: The lease disclosure rule change. *Journal of Accounting and Economics* 10 (4): 277–310.
- International Accounting Standards Committee (IASC). 1997. *Leases*. International Accounting Standard 17. London, U.K.: IASC.
- Jamal, K., and H. T. Tan. 2010. Joint effects of principles-based versus rules-based standards and auditor type in constraining financial managers' aggressive reporting. *The Accounting Review* 85 (4): 1325–1346.
- Kadous, K., S. Kennedy, and M. Peecher. 2003. The effect of quality assessment and directional goal commitment on auditors' acceptance of client-preferred accounting methods. *The Accounting Review* 78 (3): 759–778.
- Klein, A. 2002. Economics determinants of audit committee independence. *The Accounting Review* 77 (2): 435–452.
- Lander, G. P. 2004. *What is Sarbanes-Oxley?* Boston, MA: McGraw-Hill.
- Libby, R., R. Bloomfield, and M. W. Nelson. 2002. Experimental research in financial accounting. *Accounting, Organizations and Society* 27 (8): 775–810.
- Lindman, H. R. 1974. *ANOVA in Complex Experimental Designs*. San Francisco, CA: WH Freeman.
- MacKinnon, D., and J. Dwyer. 1993. Estimating mediated effects in prevention studies. *Evaluation Review* 17 (April): 144–158.
- Maines, L. A., E. Bartov, P. Fairfield, D. E. Hirst, T. E. Iannaconi, R. Mallett, C. M. Schrand, D. J. Skinner, and L. Vincent. 2003. Evaluating concepts-based vs. rules-based approaches to standard setting. *Accounting Horizons* 17 (1): 73–89.
- . 2007. Spotlight on principles-based financial reporting. *Business Horizons* 50 (5): 359–364.

- Nelson, M. W., J. A. Elliott, and R. L. Tarpley. 2002. Evidence from auditors about managers' and auditor's earnings management decisions. *The Accounting Review* 77 (Supplement): 175–202.
- . 2003. Behavioral evidence on the effects of principles- and rules-based standards. *Accounting Horizons* 17 (1): 91–104.
- Ng, T. B., and H.-T. Tan. 2003. Effects of authoritative guidance availability and audit committee effectiveness on auditors' judgments in an auditor-client negotiation context. *The Accounting Review* 78 (3): 801–818.
- Peasnell, K. V., P. Pope, and S. Young. 2000. Accrual management to meet earnings targets: UK evidence pre- and post-Cadbury. *The British Accounting Review* 32 (4): 415–445.
- Psaros, J., and K. T. Trotman. 2004. The impact of the type of accounting standards on preparers' judgments. *Abacus* 40 (1): 76–93.
- Pulliam, S. 1988. Beating FAS 13. *Corporate Finance* (December): 31.
- Sanchez, M. H., C. P. Agoglia, and R. C. Hatfield. 2007. The effect of auditors' use of a reciprocity-based strategy on auditor-client negotiations. *The Accounting Review* 82 (1): 241–263.
- Schipper, K. 2003. Principles-based accounting standards. *Accounting Horizons* 17 (1): 61–72.
- Securities and Exchange Commission (SEC). 2003. *Study Pursuant to Section 103(d) of the Sarbanes Oxley Act of 2002 on the Adoption by the United States Financial Reporting System of a Principles-Based Accounting System*. Washington, D.C.: SEC.
- . 2008. *Roadmap for the Potential Use of Financial Statements Prepared in Accordance With International Financial Reporting Standards by U.S. Issuers*. Washington, D.C.: SEC.
- . 2010. *Commission Statement in Support of Convergence and Global Accounting Standards*. Washington, D.C.: SEC.
- Shortridge, R. T., and M. Myring. 2004. Defining principles-based accounting standards. *The CPA Journal Online*, available at: www.nysscpa.org/cpajournal/2004/804/essentials/p34.htm.
- Stewart, J., and L. Munro. 2007. The impact of audit committee existence and audit committee meeting frequency on the external audit: Perceptions of Australian auditors. *International Journal of Auditing* 11 (1): 51–69.
- Trompeter, G. 1994. The effect of partner compensation schemes and generally accepted accounting principles on audit partner judgment. *Auditing: A Journal of Practice & Theory* 13 (2): 56–68.
- U.S. House of Representatives. 2002. *The Sarbanes-Oxley Act of 2002*. Public Law No. 107–204 [H.R. 3763]. Washington, D.C.: Government Printing Office.
- Weil, R. 2002. *Fundamental Causes of the Accounting Debacle at Enron: Show Me Where It Says I Can't*. Washington, D.C.: The Committee on Energy and Commerce, U.S. House of Representatives.
- Xie, B., W. N. Davidson, and P. J. DaDalt. 2003. Earnings management and corporate governance: The role of the board and the audit committee. *Journal of Corporate Finance* 9: 295–316.

Effects of SFAS 133 on the Risk Relevance of Accounting Measures of Banks' Derivative Exposures

Anwer S. Ahmed
Texas A&M University

Emre Kilic
Gerald J. Lobo
University of Houston

ABSTRACT: We provide evidence on the effects of SFAS 133 on the risk relevance of accounting measures of bank derivative exposures to bond markets. First, we find that interest rate derivatives classified as hedging are more *negatively* associated with fixed-rate bond spreads after SFAS 133. We also find that hedging derivatives offset non-trading positions to a greater extent after SFAS 133. Second, for the largest 25 banks, we find that interest and foreign exchange rate trading derivatives are more *negatively* associated with fixed-rate bond spreads after SFAS 133, consistent with more economic hedges being classified as trading after SFAS 133. For these banks, trading derivative exposures offset non-derivative trading exposures to a greater extent after SFAS 133. Our results suggest that, contrary to critics' claims, SFAS 133 has increased the risk relevance of accounting measures of derivative exposures to bond investors and benefited banks in terms of reducing their cost of capital.

Keywords: *risk relevance; derivatives; SFAS 133; bond spread.*

JEL Classifications: *G12; G14; G21; G32; M41.*

I. INTRODUCTION

We provide evidence on the effects of the Statement of Financial Accounting Standards (SFAS) 133, *Accounting for Derivative Instruments and Hedging Activities* (FASB 1998), on the risk relevance of accounting measures of bank derivative

We gratefully acknowledge comments from Mary Lea McAnally, David Harris, Tom Lys, Douglas Stevens, Martin Wu, Yildirim Yildirim, and workshop participants at the University of Illinois–Urbana Champaign, Georgia State University, Syracuse University, Florida State University, and Virginia Commonwealth University.

Editor's note: Accepted by Steven Kachelmeier.

Submitted: July 2009
Accepted: September 2010
Published Online: April 2011

exposures.¹ We focus on banks because they extensively use derivatives in both hedging and trading activities. We assess risk relevance by examining the extent to which bank fixed-rate bond spreads (hereafter, bond spreads) are associated with measures of exposures from derivatives classified as hedging (hereafter, hedging derivatives) and derivatives classified as trading (hereafter, trading derivatives). We also provide corroborating evidence on how SFAS 133 affected the risk-reducing (or risk-enhancing) role of derivatives by studying the association between changes in fair values of derivatives and changes in fair values of non-derivative positions.

These research questions are important for at least three reasons. First, even though accounting guidance for derivatives prior to SFAS 133 was inconsistent, incomplete, and lacking in transparency, there was considerable controversy over the potential benefits and costs of SFAS 133, especially for banks (see Section II for more discussion of the accounting guidance before and after SFAS 133). In particular, proponents argued that, because SFAS 133 provides greater uniformity in accounting for derivatives and standardizes the criteria for qualifying for hedge accounting, it improves the transparency and reliability of derivatives reporting by banks as well as their monitoring of risk-management strategies involving derivatives, thereby reducing their cost of capital.² However, critics argued that SFAS 133 makes it more difficult to obtain hedge accounting treatment and exacerbates the difference between accounting and economic hedges.³ Thus, to the extent banks record economic hedges as if they are speculative and investors penalize banks for these positions (Morgan and Stiroh 2001), SFAS 133 would increase banks' cost of capital. Similar views were expressed by the Financial Executives Institute (1997, 2), which stated, "[SFAS 133] will send a highly misleading message to readers of financial statements, unfavorably affecting the cost of capital." Because one of the precepts the Financial Accounting Standards Board (FASB) follows when issuing a standard is that "the costs imposed to meet the standard should be justified in relation to the overall benefits of the resulting information" (SFAS 133, ¶ 232), documenting evidence on the effects of SFAS 133 is necessary for a complete cost-benefit analysis of this standard.⁴

Second, prior accounting studies focus mostly on the value relevance of bank derivative disclosures (Barth et al. 1996; Eccher et al. 1996; Nelson 1996; Venkatachalam 1996; Ahmed et al. 2006), risk relevance of bank income measures (Beaver et al. 1970; Hodder et al. 2006), and risk relevance of disclosures to equity investors (McAnally 1996; Schrand 1997). To our knowledge, the risk relevance of accounting measures of derivative exposures to bond investors has not been examined previously. This is an important gap in the literature because (1) bonds are a much larger source of capital than stocks,⁵ and (2) there are important structural differences between bond and

¹ SFAS 138 (FASB 2000), Accounting for Certain Derivative Instruments and Certain Hedging Activities—An amendment of SFAS 133, and SFAS 149 (FASB 2003), Amendment of SFAS 133 on Derivative Instruments and Hedging Activities, which were issued later, address certain issues related to the implementation of SFAS 133, but do not change the basic model of SFAS 133.

² For example, the Investor Protection Department of the Consumer Federation of America (1997, 1) states that: "[accounting standards for derivatives and hedging activities *before* SFAS 133] leave investors in the dark about billions of dollars of transactions that the companies they invest in engage in each year. [SFAS 133] will aid investors by increasing the visibility, comparability, and understandability of the risks and benefits associated with these often complex financial instruments."

³ For example, Canadian Bankers Association (1997, 2) states that "The hedging criteria in [SFAS 133] are overly restrictive and disallow hedge accounting in situations where economic exposures are currently being effectively hedged."

⁴ The responses to the Exposure Draft of SFAS 133 indicate that banks generally opposed SFAS 133 from a cost-benefit standpoint. For example, First Chicago NBD Corporation (1997, 13–14) states: "The standard cannot be justified from a cost-benefit standpoint ... [Proposed changes] are not critical enough to compensate for the cost of implementing the standard."

⁵ Flow of Funds data published by the Board of Governors of the Federal Reserve System indicate that the size of the debt market at the end of 2005 was about twice that of the equity market.

stock markets. For example, bond markets are dominated by more sophisticated participants than are stock markets (Bhojraj and Swaminathan 2009), but are less liquid than stock markets. This study jointly investigates the risk relevance of accounting measures of derivative exposures in bond markets and how SFAS 133 altered that relevance.

Third, studying the effects of SFAS 133 on bond spreads is a more direct and cleaner way of studying the effects of SFAS 133 on the risk relevance of derivatives reporting because the primary driver of bond prices is risk, whereas stock prices are driven by changes in expected future cash flows as well as changes in risk (or discount rates).

If derivatives are used for risk reduction, then derivative exposures would be negatively associated with bond spreads. We hypothesize that derivatives classified as hedging are more negatively associated with bond spreads after SFAS 133 because SFAS 133 uses more stringent criteria for classifying derivatives as hedging derivatives and requires more detailed documentation of the hedging relationship than prior standards, thus increasing the likelihood that hedging derivatives consist of highly effective derivative hedges. On the other hand, trading derivatives will be positively (negatively) associated with bond spreads if they increase (reduce) net exposures. However, because the relative weight of derivatives used for economic hedges within the trading derivatives category is likely to increase after SFAS 133, we also hypothesize that trading derivatives will be less positively (or more negatively) associated with bond spreads after SFAS 133 if bondholders correctly perceive the increase in the relative weight of economic hedges within the trading derivatives category. We expand on these arguments in Section II.

We examine the association between bank bond spreads and accounting measures of derivative exposures before and after SFAS 133 while controlling for bond and bank characteristics identified in prior studies. Our tests are based on a sample of subordinated bank bonds with U.S. dollar denominations. We find that, after SFAS 133, exposures from interest rate hedging derivatives are more negatively associated with bond spreads. In terms of economic magnitudes, before SFAS 133, the tax-adjusted bond spread is approximately 19.5 basis points lower for a bank with the mean level of exposure from hedging derivatives relative to a bank with no such exposure, whereas it is 30.8 basis points lower after the adoption of SFAS 133. Furthermore, we find that exposures from interest and foreign exchange rate trading derivatives are also more negatively associated with bond spreads after SFAS 133 for the largest banks, which are extensively involved in derivatives trading. For these banks, the mean level of exposure from trading derivatives is associated with a 5.8 basis point lower spread before, and a 9.9 basis point lower spread after SFAS 133, compared to banks without such exposure.

To provide corroborating evidence on the risk-reducing role of derivatives, we investigate how SFAS 133 altered the relation between fair values of derivative and non-derivative exposures. Specifically, we hypothesize that, after SFAS 133, (1) hedging derivatives offset exposures from non-trading positions to a greater extent, and (2) trading derivatives offset exposures from non-derivative trading positions to a greater extent. Consistent with these hypotheses, we find that fair value changes in hedging derivatives and trading derivatives are more negatively related to fair value changes in non-trading positions and non-derivative trading positions, respectively, after SFAS 133. These results suggest that hedging derivatives offset exposures from non-trading positions, whereas trading derivatives offset exposures from non-derivative trading positions.

We contribute to the prior literature in two important ways. First, SFAS 133 has been one of the most controversial standards ever issued by the FASB. Contrary to critics' claims, we find that accounting measures of derivative exposures are more negatively associated with bond spreads after SFAS 133 than before SFAS 133, indicating that SFAS 133 has improved the risk relevance of derivatives reporting, heightened investors' confidence in the stated purposes of holding derivatives, and benefited banks in terms of reducing their cost of capital.

Second, prior work does not provide evidence on how bond investors price derivative exposures. We document that bond investors price both hedging and trading derivatives in a manner consistent with their use in offsetting non-derivative inherent risk exposures, and that these offsetting effects are more pronounced after SFAS 133. Our findings contrast with the results of prior studies (such as Barth et al. 1996; Eccher et al. 1996; Nelson 1996; Ahmed et al. 2006) that find that equity investors do not price derivative disclosures prior to SFAS 133, but are consistent with the greater sophistication of bond market investors.

Section II presents the background and hypotheses development. Section III describes the research design. Section IV presents the evidence and Section V the conclusion.

II. BACKGROUND AND HYPOTHESES DEVELOPMENT

Hedge Accounting for Derivatives before and after SFAS 133

Prior to SFAS 133, accounting for derivatives was primarily guided by SFAS 52 (FASB 1981), *Foreign Currency Translation*, and SFAS 80 (FASB 1984), *Accounting for Futures Contracts*, which addressed foreign currency transactions and exchange-traded futures contracts, respectively, and the Emerging Issues Task Force (EITF) Issue No. 84-36. There were several important inconsistencies in the requirements for hedge accounting treatment between SFAS 52 and SFAS 80.⁶ First, SFAS 52 required transaction-level risk reduction for derivatives to be classified as hedging, whereas SFAS 80 required enterprise-level risk reduction. Second, while SFAS 52 did not refer to correlation between changes in the values of hedging derivatives and the hedged items, SFAS 80 frequently emphasized a high correlation requirement for an effective hedge (e.g., ¶ 4 (b) and 11). Consequently, the Securities and Exchange Commission (SEC), under SFAS 80, interpreted a hedge to be effective if the cumulative change in the value of the hedging instrument was between 80 and 125 percent of the cumulative change in the value of the hedged item, but did not articulate a specific threshold for assessing hedge effectiveness under SFAS 52 or EITF 84-36 (Swad 1995; Carpenter 1996; Anson 1999; Finnerty and Grant 2002). Third, SFAS 52 did not require ongoing assessment of hedge effectiveness, whereas SFAS 80 did. Finally, SFAS 52 did not allow hedging of anticipated transactions, whereas SFAS 80 allowed such hedging.

In addition to these inconsistencies, there were also gaps in the accounting guidance. Neither SFAS 52 nor SFAS 80 addressed widely used derivatives such as interest rate swaps and options (Beier 1993; Montesi and Lucas 1996). Hedge accounting for swaps and options was applied by analogy to SFAS 52 and SFAS 80 (Gastineau et al. 2001). However, depending upon the sources of the analogy, the classification of and accounting for derivatives were *ad hoc* and often inconsistent across reporting entities (SFAS 133, ¶ 237). Some derivatives were accounted for as hedges regardless of whether they were part of a hedging strategy (SFAS 133, ¶ 235) or whether they were effective hedges.

The EITF attempted to address the accounting for interest rate swaps with EITF 84-36. However, in contrast to the guidance in SFAS 80, enterprise-level interest rate risk reduction was not required by EITF 84-36. Hence, prior to SFAS 133, the inconsistencies between the effective standards regarding what constituted a hedge, coupled with the failure of those standards to cover widely used derivatives contracts, made disclosure of and accounting for derivatives among the least standardized areas of financial reporting.

SFAS 133 was issued to provide uniformity in accounting for derivatives and standardize hedge accounting criteria.⁷ Regarding the former, the standard requires the recognition of all

⁶ Rane (1992), Montesi and Lucas (1996), Anson (1999), Gastineau et al. (2001) and the FASB (in SFAS 133) present detailed discussions of the inconsistencies between SFAS 52 and SFAS 80.

⁷ In order to increase transparency in derivatives activities, the FASB adopted SFAS 119 (FASB 1994), *Disclosure about*

derivatives at fair value and recognition of changes in the fair value of derivatives either in net income or in other comprehensive income. As for the latter, SFAS 133 standardized the criteria to qualify for hedge accounting and made what constitutes a hedge more precise. Most relevant to banks, under SFAS 133, derivatives that hedge certain types of risks (i.e., the risk of changes in the value or cash flows from specific exposures (i.e., individual assets or liabilities) are classified as hedging and qualify for hedge accounting.⁸ In addition, SFAS 133 requires formal and unambiguous documentation of the hedging relationship between the derivative and the specific exposure.⁹ Given that documentation of the relation between the hedged and the hedging exposures at the enterprise level is difficult, a shift in focus from enterprise-level risk reduction to specific-exposure risk reduction made the relation between the hedging and the hedged exposures easier to document and therefore more identifiable after SFAS 133.

While SFAS 133 does not provide a specific hedge effectiveness threshold, similar to the SEC's effectiveness interpretation under SFAS 80, practitioners interpret high effectiveness to mean that cumulative changes in the value of the hedging instrument should be between 80 and 125 percent of the inverse cumulative changes in the value or the cash flows of the hedged item. Accordingly, most macro hedging derivatives do not qualify for hedge accounting under SFAS 133 as there is no objective method of gauging the effectiveness of the hedging derivative without linkage to a single, identifiable asset or liability (SFAS 133, ¶ 447).

Hypothesis Development

Corporate risk management theory posits that hedging reduces costs associated with external financing (Mayers and Smith 1982; Smith and Stulz 1985; Dadalt et al. 2002). Banks hold hedging derivatives to offset exposures mainly to macroeconomic risk factors, particularly interest and foreign exchange rates. Therefore, such derivatives reduce a bank's systematic risk by making its cash flows less dependent on macroeconomic factors (Froot et al. 1993; Gay and Nam 1998), and reduce a bank's default risk by shielding its capital from losses due to unanticipated movements in market risk factors. Given the finding of Elton et al. (2001) that bond spreads consist mainly of systematic and default risk premia after tax premium is removed, derivative hedges likely result in lower bond spreads to the extent they are effective.

As discussed earlier, prior to SFAS 133, many derivatives were accounted for under EITF 84-36 or by analogy to SFAS 52, both of which conflicted with SFAS 80 in terms of risk reduction and did not contain an explicit hedge effectiveness criterion. Derivatives classified as "hedging" after SFAS 133 likely meet a higher hedge effectiveness threshold "on average" than do derivatives classified as hedging before SFAS 133 because SFAS 133 (1) links hedging derivatives with specific identifiable exposures, (2) standardizes the hedge effectiveness criteria, and (3) requires explicit documentation and ongoing assessment of the hedging relationship. The more stringent qualification and documentation rules imposed by SFAS 133 are likely to increase investors' confidence in the stated purpose of holding hedging derivatives and improve the monitoring of hedging activities, resulting in less uncertainty in the financial statements and lower bank bond spreads. This leads to our first hypothesis:

Derivative Financial Instruments and Fair Value of Financial Instruments. SFAS 119, which was superseded by SFAS 133, regulated only the derivative-related disclosures, but neither changed the existing accounting for derivatives nor the criteria for qualifying as a hedging derivative. Instead, it relied on the classification criteria provided by SFAS 52 and SFAS 80.

⁸ SFAS 133 also allows hedging of off-balance sheet firm commitments and forecasted transactions.

⁹ Documentation under SFAS 133 should include identification of (1) the hedging derivative, (2) the hedged asset or liability, (3) the nature of the risk being hedged, and (4) how the effectiveness of the hedge is assessed.

H1: Bank bond spreads are more negatively related to derivatives classified as hedging after SFAS 133 than before SFAS 133.

Bank trading derivatives consist of customer-related positions, proprietary positions, and economic hedges. Banks typically take offsetting positions in customer-related derivatives trading. By matching customer-related derivatives in terms of market risk exposure, a bank can generate fee income without assuming any market risk. Although not as common, some banks also take non-offset positions in customer-related derivatives trading (Puwaliski 2003). Hence, whether customer-related derivatives increase or reduce overall market risk depends on whether the residual market risk, if any, from these positions offsets or adds to other risk exposures. Unlike customer-related derivatives, proprietary trading derivatives are likely to increase exposure because they are held for speculation. Lastly, some derivatives classified as trading are in fact economic hedges of other exposures.¹⁰ Therefore, bank derivatives classified as trading include risk-reducing positions (i.e., economic hedges), risk-enhancing positions (i.e., speculative positions), and positions with ambiguous risk implications (i.e., customer-related positions). However, the relative weights of these positions are not disclosed under SFAS 119 (FASB 1994) or under SFAS 133. Thus, the relation between bond spreads and trading derivatives could be negative or positive, depending upon whether the net trading derivatives positions reduce or increase overall risk. However, the relative importance of economic hedges within trading derivatives is likely to increase after SFAS 133 for at least two reasons. First, as discussed earlier, SFAS 133 prohibits the reporting of derivatives used in macro hedging as hedging derivatives, and requires them to be classified as trading derivatives. Second, some derivatives used by banks in *de facto* hedging transactions are classified as trading because they are no longer defined as hedging derivatives, or are not sufficiently effective under SFAS 133. Such derivatives include hedges of interest rate risk in held-to-maturity securities (SFAS 133, ¶ 21 (d) and ¶ 29 (e)), long-term core deposits (SFAS 133, ¶ 317 and ¶ 437), and prepayment risks of financial assets or liabilities (SFAS 133, ¶ 29 (h)).

Banks commonly macro-hedge gap positions or value-at-risk. Therefore, they opposed the disqualification of macro hedges from hedge accounting in their response letters to the Exposure Draft for SFAS 133. The general view of the banking industry was that reporting some legitimate economic hedges as trading, as required by SFAS 133, would generate potentially misleading and unfavorable results when applied to banks (America's Community Bankers 1997). While these economic hedges are typically less effective hedges than the derivatives that qualify for hedge accounting under SFAS 133, they are more effective hedges than derivatives classified as trading prior to SFAS 133. Therefore, if creditors price trading derivatives in a manner consistent with these derivatives including more economic hedges after SFAS 133, a more negative (or less positive) association between bank bond spreads and trading derivatives will result. Thus, we test the following hypothesis:

H2: Bank bond spreads are more negatively (or less positively) related to derivatives classified as trading after SFAS 133 than before SFAS 133.

Derivatives that hedge exposures from trading positions are required to be classified as trading (SFAS 119, ¶ 47), whereas derivatives that hedge exposures from non-trading assets and liabilities, such as loans and deposits, are to be classified as hedging. One of the presumptions underlying H1 is that derivatives classified as hedging after SFAS 133 are, on average, more effective hedges than

¹⁰ For example, in its 10-K for 1999, Compass Bancshares, Inc. states: "The primary purposes for using interest rate derivatives in the trading account are to facilitate customer transactions and to *help protect cash market positions in the trading account against interest rate movement*. [Some options in trading account] assist in *protecting the fair value of the trading account* against rising short-term interest rates while maintaining limited risk to declining rates." (emphasis added)

derivatives classified as hedging before SFAS 133 because SFAS 133 restricts derivatives classified as hedging to highly effective derivative hedges of specific exposures. Therefore, we expect derivatives classified as hedging to offset the corresponding hedged items to a greater extent after SFAS 133, which leads to the following hypothesis:

H3: The relation between change in fair value of derivatives classified as hedging and change in fair value of net non-trading assets is more negative after SFAS 133 than before SFAS 133.

Unlike fair values of hedging derivatives, fair values of trading derivatives need not covary with other existing exposures in any specific direction to be classified as trading under either SFAS 119 or SFAS 133. Therefore, whether trading derivatives offset or add to the exposures from non-derivative trading or non-trading positions is an empirical question. Liu et al. (2004) do not find significant cross-hedging between trading and non-trading exposures. On the other hand, trading derivatives include derivatives that are economic hedges of other trading positions. SFAS 133 likely increases the average weight of such economic hedges in bank derivative portfolios classified as trading for two reasons. First, as mentioned earlier, some derivatives that are classified as hedging prior to SFAS 133 are classified as trading after SFAS 133. Because recognizing the hedging derivative at fair value, but the hedged position at cost would create income volatility, SFAS 133 allows banks to transfer such hedged items into the trading category along with derivatives hedging these positions (SFAS 133, ¶ 54). This likely increases the weight of derivative and non-derivative positions with offsetting exposures in the trading portfolio.¹¹ Second, banks could choose to avoid designating some derivatives as hedging and instead classify them as trading after SFAS 133 because the costs associated with the more onerous documentation and more complex hedge accounting requirements under SFAS 133 may outweigh the benefits.¹² This reasoning leads to the following hypothesis:

H4: The relation between change in fair value of derivatives classified as trading and change in fair value of net non-derivative trading assets is more negative (or less positive) after SFAS 133 than before SFAS 133.

III. RESEARCH DESIGN AND EMPIRICAL MODELS

Model for Testing Hypotheses 1 and 2

Hypotheses 1 and 2 posit that the relation between accounting measures of derivative exposures and bond spreads will differ before and after SFAS 133. We test these hypotheses with a model that relates accounting measures of derivative exposures to bond spreads, controlling for bond characteristics (Hancock and Kwast 2001),¹³ bank characteristics (Flannery and Sorescu 1996; Morgan and Stiroh 2001; Hancock and Kwast 2001; Jagtiani et al. 2002), and off-balance sheet instruments other than derivatives that are potentially used as economic hedges. We estimate the model separately for the periods before and after SFAS 133. The model is specified as follows:

$$\begin{aligned}
 r_{it} = & \beta_0 + \beta_1 \text{Big}_{it}^{25} + \beta_2 \text{IrHdg}_{it} + \beta_3 \text{Big}_{it}^{25} * \text{IrHdg}_{it} + \beta_4 \text{DeltaIrHdg}_{it} + \beta_5 \text{Big}_{it}^{25} * \text{DeltaIrHdg}_{it} \\
 & + \beta_6 \text{FxHdg}_{it} + \beta_7 \text{Big}_{it}^{25} * \text{FxHdg}_{it} + \beta_8 \text{DeltaFxHdg}_{it} + \beta_9 \text{Big}_{it}^{25} * \text{DeltaFxHdg}_{it} \\
 & + \beta_{10} \text{IrTrd}_{it} + \beta_{11} \text{Big}_{it}^{25} * \text{IrTrd}_{it} + \beta_{12} \text{DeltaIrTrd}_{it} + \beta_{13} \text{Big}_{it}^{25} * \text{DeltaIrTrd}_{it} + \beta_{14} \text{FxTrd}_{it}
 \end{aligned}$$

¹¹ Reclassification of some securities is allowed under SFAS 133 to mitigate the likely effect of the standard on earnings volatility. Recognizing both the derivatives and the hedged items at fair value after SFAS 133 allows the offsetting gains and losses from the hedging and the hedged positions to be recognized in income concurrently.

¹² In line with this view, a survey by Bank of America (2006) revealed that a small but growing number of banks opted not to apply hedge accounting to certain derivative hedges because the cost and effort of effectiveness testing and documentation required under SFAS 133 are too high.

¹³ We restrict the sample to subordinated bank bonds with U.S. dollar denominations that are all non-callable, non-putable, non-sinking fund, and non-convertible. Therefore, bonds in our sample vary in maturity and amount only.

$$\begin{aligned}
& + \beta_{15} \text{Big}_{it}^{25} * \text{FxTrd}_{it} + \beta_{16} \text{DeltaFxTrd}_{it} + \beta_{17} \text{Big}_{it}^{25} * \text{DeltaFxTrd}_{it} + \beta_{18} \text{Maturity}_{it} \\
& + \beta_{19} \text{Amount}_{it} + \beta_{20} \text{Size}_{it} + \beta_{21} \text{Insured}_{it} + \beta_{22} \text{Estate}_{it} + \beta_{23} \text{Npl}_{it} + \beta_{24} \text{Gap}_{it} \\
& + \beta_{25} \text{NonDerTrd}_{it} + \beta_{26} \text{Cap}_{it} + \beta_{27} \text{Roa}_{it} + \beta_{28} \text{CrRisk}_{it} + \beta_{29} \text{LoanComm}_{it} \\
& + \beta_{30} \text{Guarantee}_{it} + \beta_{31} \text{CrDer}_{it} + \beta_{32} \text{Strip}_{it} + \beta_{33} \text{Iml}_{it} + \varepsilon_{it}, \tag{1}
\end{aligned}$$

where for bank i at time t :

- r_{it} = tax-adjusted spread;
- Big_{it}^{25} = 1 if the bank is one of the largest 25 banks in the U.S. with respect to asset size, and 0 otherwise;
- IrHdg_{it} = absolute value of net long notional amount of interest rate derivatives classified as hedging scaled by total assets;
- FxHdg_{it} = absolute value of net long notional amount of foreign exchange derivatives classified as hedging scaled by total assets;
- DeltaIrHdg_{it} = absolute value of the annual change in the fair value of interest rate derivatives classified as hedging divided by the annual change in the interest rates scaled by total assets;
- DeltaFxHdg_{it} = absolute value of the annual change in the fair value of foreign exchange derivatives classified as hedging divided by the annual change in the foreign exchange rates scaled by total assets;
- IrTrd_{it} = absolute value of net long notional amount of interest rate derivatives classified as trading scaled by total assets;
- FxTrd_{it} = absolute value of net long notional amount of foreign exchange derivatives classified as trading scaled by total assets;
- DeltaIrTrd_{it} = absolute value of the annual change in the fair value of interest rate derivatives classified as trading divided by the annual change in the interest rates scaled by total assets;
- DeltaFxTrd_{it} = absolute value of the annual change in the fair value of foreign exchange derivatives classified as trading divided by the annual change in the foreign exchange rates scaled by total assets;
- Maturity_{it} = natural logarithm of the bond's years to maturity;
- Amount_{it} = natural logarithm of the size of bond issue in million dollars;
- Size_{it} = natural logarithm of bank's total assets;
- Insured_{it} = ratio of total insured deposits to total deposits;
- Estate_{it} = real estate acquired in satisfaction of debt previously contracted scaled by total assets;
- Npl_{it} = sum of nonaccrual loans and accruing loans past due 90 days or more scaled by total assets;
- Gap_{it} = absolute value of the difference between bank assets and liabilities maturing or repricing within one year scaled by total assets;
- Roa_{it} = annual net income scaled by total assets;
- Cap_{it} = ratio of book value of equity to risk-weighted assets;
- NonDerTrd_{it} = fair value of non-derivative trading assets scaled by total assets;
- CrRisk_{it} = bank's credit losses incurred on off-balance sheet derivative contracts scaled by total assets;
- LoanComm_{it} = amount of loans for which official promises to lend are conveyed scaled by total assets;

$Guarantee_{it}$ = on- and off-balance sheet guarantees and insurance contracts that back financial obligations scaled by total assets;

$CrDer_{it}$ = notional amount of credit derivatives on which the bank is the beneficiary scaled by total assets;

$Strip_{it}$ = fair value of principal-only strips scaled by total assets; and

Iml_{it} = Inverse Mills ratio obtained from the first-stage multinomial logit model.

We provide the list of data sources and discuss the rationale for including each of the control variables in Appendix A.

Interest rate swaps and options constitute the majority of bank derivative portfolios. Ryan (2007) indicates that nonlinearity of option payoff structure makes the assessment of options' effectiveness as hedging instruments difficult when the hedged item has a linear payoff structure. He also indicates that bank market risk disclosures, which are typically linear characterizations of exposures, cannot sufficiently reflect nonlinear exposures. Also, accounting for different types of derivatives under different pronouncements (e.g., forward under SFAS 52, interest rate swaps under EITF 84-36) resulted in differences in hedge accounting for different types of derivatives prior to SFAS 133. Therefore, the effect of SFAS 133 could vary based on the type of derivative. In additional sensitivity analysis, we examine whether a particular type of interest rate derivative drives our results. We re-estimate the model presented in Equation (1) only for interest rate derivatives while allowing the coefficients on (linear) swaps and (non-linear) options to vary.

Since most of the derivatives activity in the U.S. is dominated by the largest 25 banks, we allow the coefficients on the derivative exposure variables to vary across the largest 25 banks and the remaining banks in order to control for the potential difference in risk relevance of information on derivative exposures for these two groups of banks. As hypothesized in H1, we expect more negative coefficients on the accounting measures of exposures from derivatives classified as hedging (i.e., $\beta_2 + \beta_3$, $\beta_4 + \beta_5$, $\beta_6 + \beta_7$, and $\beta_8 + \beta_9$ for the largest 25 and β_2 , β_4 , β_6 , and β_8 for the remaining banks) after the adoption of SFAS 133. Also, as discussed before, whether trading derivatives are positively or negatively related to bond spreads is an empirical question. However, as hypothesized in H2, we expect the coefficients on the accounting measures of exposures from derivatives classified as trading (i.e., $\beta_{10} + \beta_{11}$, $\beta_{12} + \beta_{13}$, $\beta_{14} + \beta_{15}$, and $\beta_{16} + \beta_{17}$ for the largest 25, and β_{10} , β_{12} , β_{14} , and β_{16} for the remaining banks) to be more negative (or less positive) after SFAS 133.

Given our research questions, our study naturally focuses on the subsample of banks that use derivatives. To the extent that the characteristics of derivative-user banks systematically differ from those of non-derivative-user banks, an empirical model that ignores the non-randomness of this subsample likely suffers from selection bias. We control for the potential effects of selection bias by using a standard extension of the Heckman two-stage procedure generalized to polychotomous choices developed by Lee (1983), where the bond-pricing model presented in Equation (1) constitutes the second stage. In the first stage, we classify banks into three categories: banks without derivatives (*Nonuser*), banks with only hedging derivatives (*Hedger*), and banks with both hedging and trading derivatives (*Trader*). Because *Trader* banks are typically those that engage in derivatives activities extensively, the last two categories jointly control for the extent of derivative use and the type of activity (i.e., hedging versus trading). We include the Inverse Mills ratios (*Iml*) obtained from the multinomial logit model in the bond-pricing model.

Finance theory posits that banks with smaller asset portfolio size (*Size*), lower leverage (*Lev*), lower interest rate risk (*Gap*), lower non-performing loans (*Npl*), less convex tax functions (*Convexity*), higher net interest margin (*Nim*), and stronger managerial incentives to take risk (*Incentive*) are less likely to use derivatives for hedging purposes. Prior finance studies also suggest that the likelihood of banks to use derivatives for trading purposes increases with *Size*, liquidity

(*Liquidity*), the size of the commercial and industrial loan portfolio relative to total assets (*ComIndLoan*), and decreases with *Lev* and *Nim*.¹⁴ We specify *Hedger* as the base category in the following first-stage multinomial logit model:

$$D_{it}(\text{Nonuser/Hedger}) = \alpha_0 + \alpha_1 \text{Big}_{it}^{25} + \alpha_2 \text{Size}_{it} + \alpha_3 \text{Lev}_{it} + \alpha_4 \text{Gap}_{it} + \alpha_5 \text{Nim}_{it} + \alpha_6 \text{Npl}_{it} \\ + \alpha_7 \text{Liquidity}_{it} + \alpha_8 \text{Incentive}_{it} + \alpha_9 \text{Convexity}_{it} + \alpha_{10} \text{ComIndLoan}_{it} \\ + \varepsilon_{it} \quad (2a)$$

$$D_{it}(\text{Trader/Hedger}) = \theta_0 + \theta_1 \text{Big}_{it}^3 + \theta_2 \text{Big}_{it}^{4-25} + \theta_3 \text{Size}_{it} + \theta_4 \text{Lev}_{it} + \theta_5 \text{Gap}_{it} + \theta_6 \text{Nim}_{it} + \theta_7 \text{Npl}_{it} \\ + \theta_8 \text{Liquidity}_{it} + \theta_9 \text{Incentive}_{it} + \theta_{10} \text{Convexity}_{it} + \theta_{11} \text{ComIndLoan}_{it} + \varepsilon_{it}, \quad (2b)$$

where, for bank i at time t , D_{it} equals 0 if bank i is a *Hedger* in both Equations (2a) and (2b), equals 1 in Equation (2a) if bank i is a *Nonuser*, and equals 1 in Equation (2b) if bank i is a *Trader*. Big^{25} equals 1 if the observation belongs to one of the largest 25 banks in terms of asset size, and 0 otherwise in Equation (2a). Big^3 (Big^{4-25}) equals 1 if the observation belongs to one of the largest three (four to 25) banks in terms of asset size, and 0 otherwise in Equation (2b).¹⁵ We discuss the rationale for each of the variables in Appendix B.

Convexity serves as an exclusion restriction in our two-stage estimation procedure because, while both finance theory and empirical evidence suggest a direct link between tax convexity and derivative usage (Nance et al. 1993; Tufano 1996; Géczy et al. 1997; Graham and Rogers 2002), it is unlikely that tax convexity and bond spread are directly linked.¹⁶

Measurement of Bond Yield Spread

Following Elton et al. (2001), we use a tax-adjusted measure of bond yield spread that eliminates distortions stemming from mismatches of duration and convexity between bank bonds and benchmark Treasury bonds.¹⁷ We present the details of this technique in Appendix C.

Measurement of Derivative Exposure

Ideally, a measure of derivative exposure disaggregates notional and fair values by long and short positions taken, class of instrument, time to maturity, degree of leverage, and level of complexity (Hull 1993; Wong 2000). However, neither SFAS 119 nor SFAS 133 requires disclosures detailed enough to obtain an ideal measure of derivative exposure. Wong (2000) suggests a comprehensive measure of derivative exposure within the constraints of the limited available disclosures. This measure includes the size of the portfolio exposed to market risk (measured as the notional value of long derivatives net of short derivatives) and the delta of the portfolio (i.e., the rate of change of the portfolio value with respect to a change in the underlying risk factor) because derivative exposure is driven by both these variables.

¹⁴ Relevant studies from the finance literature are cited in Appendix B in the discussion of variables used in the multinomial logit model.

¹⁵ Equation (2a) does not include separate, binary size variables for the largest three and largest four to 25 banks because none of the largest three banks is in the *Nonuser* or *Hedger* category.

¹⁶ Imposing an exclusion restriction in the Heckman procedure is critical because the absence of such a restriction can lead to severe multicollinearity problems in the second-stage estimation (Francis et al. 2010).

¹⁷ The tax and accounting literatures identify various settings where taxes are impounded in the prices of assets with tax-favored (or tax-disfavored) returns (Shackelford and Shevlin 2001). Holders of corporate bonds pay state and local taxes on interest payments, whereas holders of Treasury bonds do not. Therefore, the asymmetric tax treatment of corporate and Treasury bond interest payments likely results in corporate bonds offering higher pre-tax return to yield the same after-tax return as Treasury bonds.

Our measures of derivative exposure, which account for the complementary relation among notional amounts, fair values, and past changes in the underlying risk factors (e.g., interest rates), are based on Wong (2000). For hedging derivatives, we measure portfolio size as the net long notional amount (*IrHdg* and *FxHdg*) because banks typically disclose long and short positions in such derivatives when providing information about their hedging activities.¹⁸ However, more than half the sample banks with trading derivatives do not disclose long and short positions in such derivatives separately. For these banks, we estimate the extent of the net exposure to market risk using the technique employed by the Office of the Comptroller of the Currency.¹⁹ We compute the deltas of the portfolio for each group of derivatives (*DeltaIrHdg*, *DeltaFxHdg*, *DeltaIrTrd*, and *DeltaFxTrd*) as the change in fair value of the derivative portfolio divided by the change in the underlying market risk factor (interest rate or foreign exchange rate) during the fiscal year.^{20,21}

Models for Testing Hypotheses 3 and 4

The bond-pricing models provide evidence on how bond investors perceive and price accounting measures of derivative exposures before and after SFAS 133, but do not directly test whether and how SFAS 133 altered the risk-offsetting relation between derivative and non-derivative exposures. Studies that examine how derivative exposures covary with other exposures typically examine the association between fair value gains/losses from derivatives and fair value gains/losses from non-derivative exposures (e.g., Venkatachalam 1996; Liu et al. 2004). Following Liu et al. (2004), we (1) regress the quarterly change in the net fair value of hedging derivatives ($\Delta DerHedging_{it}$) on the quarterly change in fair value of net non-derivative trading assets ($\Delta NonDerTrading_{it}$), and (2) regress the quarterly change in the net fair value of trading derivatives ($\Delta DerTrading_{it}$) on the quarterly change in fair value of net non-trading assets ($\Delta NonTrading_{it}$), respectively. A negative (positive) association of the dependent variable with $\Delta NonDerTrading$ and $\Delta NonTrading$ suggests that derivatives hedge (add to) the risks from non-derivative trading and non-trading exposures, respectively. We estimate the following models for the periods before and after SFAS 133:

$$\Delta DerHedging_{it} = \gamma_0 + \gamma_1 High_{it}^{Hdg} + \gamma_2 \Delta NonDerTrading_{it} + \gamma_3 High_{it}^{Hdg} * \Delta NonDerTrading_{it} + \gamma_4 \Delta NonTrading_{it} + \gamma_5 High_{it}^{Hdg} * \Delta NonTrading_{it} + e_{it}. \quad (3)$$

$$\Delta DerTrading_{it} = \lambda_0 + \lambda_1 High_{it}^{Trd} + \lambda_2 \Delta NonDerTrading_{it} + \lambda_3 High_{it}^{Trd} * \Delta NonDerTrading_{it} + \lambda_4 \Delta NonTrading_{it} + \lambda_5 High_{it}^{Trd} * \Delta NonTrading_{it} + e_{it}. \quad (4)$$

where $High_{it}^{Hdg}$ ($High_{it}^{Trd}$) equals 1 if the notional amount of derivatives classified as hedging (trading) scaled by beginning book value of equity is above the sample median for the year, and 0 otherwise. All other variables are scaled by beginning book value of equity. We interact $High_{it}^{Hdg}$ and $High_{it}^{Trd}$ with the fair value changes in non-derivative positions to control for potential differences in the covariation of derivative and non-derivative exposures across banks that hedge more

¹⁸ For the few sample banks that hold basis swaps, we add the notional values of the basis swaps to the absolute difference between the notional values of long and short derivatives.

¹⁹ In its derivative reports, the Office of the Comptroller of the Currency estimates the extent of non-offset derivative exposure to risk by dividing the absolute value of the net fair value of the derivative portfolio by the absolute value of the gross fair value of the derivative portfolio that is exposed to that specific risk.

²⁰ The terms of the interest rates used in the calculation of interest rate derivative portfolio deltas are based on the weighted average maturity of interest rate derivatives estimated using Call Report data. Then we compute the yield on a synthetic Treasury bond with the same maturity using the parametric yield curve specification suggested by Gurkaynak et al. (2006).

²¹ The changes in foreign exchange rates are based on the Nominal Broad Dollar Index released by the Federal Reserve.

versus less and banks that trade more versus less, respectively, because banks with larger portfolios of derivatives are more likely to be impacted by SFAS 133. As predicted in H3, we expect the coefficient on $\Delta NonDerTrading$ in Equation (3) (i.e., γ_2 for banks with less and $\gamma_2 + \gamma_3$ for banks with more hedging derivatives) to be more negative after SFAS 133 because we expect hedging derivatives to comprise more effective hedges on average after SFAS 133. We also expect the coefficient on $\Delta NonTrading$ in Equation (4) (i.e., λ_4 for banks with less and $\lambda_4 + \lambda_5$ for banks with more trading derivatives) to be more negative (or less positive) as predicted in H4 because we expect SFAS 133 to increase the weight of derivative and non-derivative positions with offsetting exposures in the trading portfolio.

IV. EVIDENCE

Sample Selection

The primary criteria for inclusion in the sample are use of derivative instruments and availability of bond issue data. To ensure that financial data are available at the time of bond issuance, we match the Call Report (FR Y-9C) and 10-K data with the data for the first subordinated bond issue after the date of the 10-K filing. Consistent with the prior bond-pricing literature, our bond sample consists of subordinated bonds because subordinated bond spreads' greater sensitivity to credit and market risks facilitates more powerful tests of the relation between bank bond spread and accounting risk measures. We obtain bond issue data from the Mergent Fixed Income Securities Database (FISD).

We restrict our sample period to two years before and two years after SFAS 133 in order to mitigate the possibility that our results are driven by factors other than SFAS 133.²² Using the Bank Compustat Database, we first identify 710 banks including both national commercial banks (SIC 6021) and state commercial banks (SIC 6022) that were publicly traded between 1999 and 2003. Using the Federal Reserve Bank of Chicago's Bank Regulatory Database, we identify 165 of these 710 banks as derivative users. We eliminate 24 of the 165 user banks due to lack of derivative disclosures in their 10-Ks, resulting in a final sample of 141 banks. We retain only one subordinated bond issue per bank in each year. These criteria yield a sample of 270 bond issues for the pre-SFAS 133 period and 265 bond issues for the post-SFAS 133 period.

Descriptive Statistics

We present descriptive statistics on the composition of sample bank derivative portfolios based on gross notional amounts in Panels A through D of Table 1. All 141 sample banks hold hedging derivatives, whereas only 52 (67) banks hold trading derivatives before (after) SFAS 133. As expected, banks carry mostly interest rate and foreign exchange derivatives. As presented in Panel A (Panel C), interest rate derivatives constitute the largest portion of derivatives, comprising 92.5 (67.4) percent and 88.3 (73.9) percent of hedging (trading) derivatives before and after SFAS 133, respectively. The weight of foreign exchange derivative contracts is higher in trading derivatives compared to hedging derivatives because the former include derivatives sold to customers that manage foreign exchange risk exposure. Swaps are the primary component of interest rate derivatives classified as hedging and as trading, whereas forwards are most prevalent among foreign exchange derivatives (Panels B and D, respectively).

Table 2 reports descriptive statistics for the dependent and independent variables employed in the bond-pricing tests. Tax-adjusted bond spread has a sample mean (median) of 70 (62) basis

²² The tests for the pre-SFAS 133 (post-SFAS 133) period are based on financial statement data for fiscal years ending December 31, 1999 and 2000 (2002 and 2003). We exclude the year 2001, the first effective year of SFAS 133, because the annual change in fair values of hedging and trading derivatives from year 2000 to 2001 is distorted by reclassification of some derivatives into trading and hedging by SFAS 133.

TABLE 1
Composition of Derivative Portfolios Based on Gross Notional Amounts

Panel A: Composition of Hedging Derivatives Based on the Type of Underlying Risk Factor

	Pre-SFAS 133 (1999–2000)			Post-SFAS 133 (2002–2003)		
	No. of User Banks	Mean	Median	No. of User Banks	Mean	Median
Interest Rate Derivatives	139	0.925	1.000	139	0.883	1.000
Foreign Exchange Derivatives	26	0.057	0.000	22	0.065	0.000
Equity Derivatives	15	0.018	0.000	19	0.052	0.000
Commodity Derivatives	3	0.001	0.000	7	0.001	0.000

Panel B: Composition of Hedging Derivatives Based on the Type of Instrument

	Pre-SFAS 133 (1999–2000)			Post-SFAS 133 (2002–2003)		
	No. of User Banks	Mean	Median	No. of User Banks	Mean	Median
Interest Rate Derivatives						
Swaps	135	0.707	0.726	131	0.712	0.755
Options (Caps and Floors)	72	0.182	0.256	80	0.202	0.188
Forwards	18	0.036	0.000	11	0.013	0.000
Foreign Exchange Derivatives						
Swaps	5	0.001	0.000	2	0.000	0.000
Options	11	0.006	0.000	8	0.012	0.000
Forwards	19	0.050	0.000	16	0.009	0.000

Panel C: Composition of Trading Derivatives Based on the Type of Underlying Risk Factor

	Pre-SFAS 133 (1999–2000)			Post-SFAS 133 (2002–2003)		
	No. of User Banks	Mean	Median	No. of User Banks	Mean	Median
Interest Rate Derivatives	51	0.674	0.844	65	0.739	0.902
Foreign Exchange Derivatives	38	0.280	0.102	39	0.186	0.033

(continued on next page)

Panel C: Composition of Trading Derivatives Based on the Type of Underlying Risk Factor

	Pre-SFAS 133 (1999–2000)			Post-SFAS 133 (2002–2003)		
	No. of User Banks	Mean	Median	No. of User Banks	Mean	Median
Equity Derivatives	17	0.041	0.000	19	0.058	0.000
Commodity Derivatives	5	0.005	0.000	8	0.018	0.000

Panel D: Composition of Trading Derivatives Based on the Type of Instrument

	Pre-SFAS 133 (1999–2000)			Post-SFAS 133 (2002–2003)		
	No. of User Banks	Mean	Median	No. of User Banks	Mean	Median
Interest Rate Derivatives						
Swaps	50	0.476	0.367	54	0.494	0.408
Options (Caps and Floors)	30	0.152	0.098	35	0.178	0.102
Forwards	7	0.047	0.000	13	0.059	0.000
Foreign Exchange Derivatives						
Swaps	6	0.022	0.000	8	0.001	0.000
Options	19	0.048	0.000	26	0.031	0.000
Forwards	32	0.206	0.182	28	0.162	0.147

Statistics presented in Panels A and B are based on 141 banks with hedging derivatives.

Pre-SFAS 133 and post-SFAS 133 statistics presented in Panels C and D are based on 52 and 67 banks with trading derivatives, respectively.

Data presented in Panels A through D are as of December 31 of each year.

TABLE 2
Descriptive Statistics for Variables Employed in Bond-Pricing Tests

Variables	Mean	Median	Minimum	Maximum
Dependent Variable				
<i>Tax-Adjusted Spread (%)</i>	0.702	0.624	0.169	8.482
Derivative Exposure Variables Based on Portfolio Size				
<i>IrHdg</i>	0.085	0.079	0.005	0.603
<i>FxHdg</i>	0.001	0.000	0.000	0.068
<i>IrTrd^a</i>	0.045	0.030	0.003	0.716
<i>FxTrd^a</i>	0.012	0.002	0.000	0.452
Derivative Exposure Variables Based on Portfolio Delta				
<i>DeltaIrHdg</i>	0.004	0.004	0.000	0.025
<i>DeltaFxHdg</i>	0.001	0.000	0.000	0.004
<i>DeltaIrTrd^a</i>	0.001	0.000	0.000	0.017
<i>DeltaFxTrd^a</i>	0.001	0.000	0.000	0.010
Control Variables				
<i>Maturity</i>	11.785	10.000	1.000	30.000
<i>Amount</i>	17.443	17.076	6.451	24.635
<i>Size</i>	15.901	15.844	12.218	20.447
<i>Insured</i>	0.411	0.479	0.232	0.867
<i>Estate</i>	0.001	0.000	0.000	0.015
<i>Npl</i>	0.007	0.006	0.000	0.057
<i>Gap</i>	0.181	0.152	0.000	0.689
<i>Roa</i>	0.012	0.011	-0.061	0.077
<i>Cap</i>	0.076	0.078	0.033	0.237
<i>NonDerTrd</i>	0.042	0.019	0.000	0.266
<i>CrRisk</i>	0.000	0.000	0.000	0.001
<i>LoanComm</i>	0.064	0.056	0.000	0.465
<i>Guarantee</i>	0.007	0.000	0.000	0.153
<i>CrDer</i>	0.005	0.000	0.000	0.254
<i>Strip</i>	0.001	0.000	0.000	0.021

The sample represents 535 bank-year observations from 1999 to 2003 (excluding 2001) for 141 banks.

All control variables are as defined in Appendix A.

^a Statistics presented are based only on banks with trading derivatives.

Variable Definitions:

IrHdg (FxHdg) = absolute value of net long/short notional value of interest rate (foreign exchange) hedging derivatives scaled by total assets;

IrTrd (FxTrd) = absolute value of net long/short notional value of interest rate (foreign exchange) trading derivatives scaled by total assets;

DeltaIrHdg (DeltaFxHdg) = portfolio delta of interest rate (foreign exchange) hedging derivatives; and

DeltaIrTrd (DeltaFxTrd) = portfolio delta of interest rate (foreign exchange) trading derivatives.

points. The mean ratio of *net* notional amount of interest rate hedging derivatives to total assets (*IrHdg*) is 8.5 percent, whereas the mean ratio of *gross* notional amount of interest rate hedging derivatives to total assets (unreported) is 12.8 percent. For banks with trading derivatives, the mean ratio of *net* notional amount of interest rate trading derivatives to total assets (*IrTrd*) is 4.5 percent, whereas the mean ratio of *gross* notional amount of interest rate trading derivatives to total assets (unreported) is 163.4 percent. These statistics indicate that, although the volume of trading derivatives is significantly greater than the volume of hedging derivatives, the size of net

exposure from hedging derivatives (*IrHdg*) is greater, on average, than the size of net exposure from trading derivatives (*IrTrd*), suggesting that banks take offsetting positions in trading derivatives to a great extent.

Unreported correlation analyses generally yield insignificant correlations between tax-adjusted spread and the variables of interest. One exception is *IrHdg*, which is negatively correlated with tax-adjusted spread both before ($r = -0.102$; $p = 0.094$) and after SFAS 133 ($r = -0.123$; $p = 0.046$). Also, the insignificant correlations (unreported) among the variables of interest suggest that multicollinearity between these variables of interest is unlikely to significantly affect our estimation results.

Results of the First-Stage Multinomial Logit Model

Table 3 contains results of the multinomial logit regressions for the *Nonuser* and the *Trader* categories. The results indicate that *Nonuser*, *Hedger*, and *Trader* banks systematically differ. Specifically, the results of the multinomial logit regression for *Nonuser* status indicate that banks less likely to use derivatives have a smaller asset portfolio (*Size*), lower probability of financial distress (*Lev*), lower interest rate risk exposure (*Gap*), lower credit risk (*Npl*), executive compensation plans with stronger incentives to take risk (*Incentive*), and less convex tax functions (*Convexity*), relative to *Hedger* banks. The results of the multinomial logit regression for *Trader* status indicate that banks more likely to use derivatives to a large extent and engage in derivatives trading have a larger asset portfolio (*Size*), lower probability of financial distress (*Lev*), lower net interest margin (*Nim*), and larger portfolio of commercial and industrial loans (*ComIndLoan*),

TABLE 3
Multinomial Logit Results

	Nonuser/Hedger			Trader/Hedger		
	Predicted Sign	Coefficient	p-value	Predicted Sign	Coefficient	p-value
<i>Big</i> ²⁵	—	−0.099	0.133			
<i>Big</i> ³				+	0.211	0.008
<i>Big</i> ^{4–25}				+	0.135	0.002
<i>Size</i>	—	−0.418	<0.001	+	0.365	0.004
<i>Lev</i>	—	−0.007	0.003	—	−0.005	0.025
<i>Gap</i>	—	−0.400	0.016	−/+	0.246	0.240
<i>Nim</i>	−/+	−3.079	0.007	—	−1.911	0.060
<i>Npl</i>	—	−1.284	0.002	−/+	0.921	0.084
<i>Liquidity</i>	−/+	0.609	0.247	+	0.035	0.162
<i>Incentive</i>	+	0.048	<0.001	−/+	0.033	0.181
<i>Convexity</i>	—	−0.005	0.053	−/+	0.002	0.014
<i>ComIndLoan</i>	−/+	0.152	0.042	+	0.542	0.072
Log Likelihood			−196.224			−368.357

The sample represents observations from 1999 to 2003 (excluding 2001).

The sample includes 710 publicly traded banks, 141 of which are derivative users. The tests are based on 10,920 bank-quarters: 8,720 observations belong to *Nonuser* banks, 1,268 observations belong to *Hedger* banks, and 932 observations belong to *Trader* banks.

All variables are as defined in Appendix B.

relative to *Hedger* banks. These findings are in line with risk management theory and the findings of prior empirical studies.²³

Results of Tests of Hypotheses 1 and 2

The data set used in this study contains observations on multiple banks across multiple years. To estimate standard errors that are robust to both time-series and cross-sectional correlation, all results are based on standard errors clustered simultaneously by bank and by year.

Table 4, Panel A reports estimation results for the model presented in Equation (1). The coefficient signs on the control variables are generally consistent with those reported in prior studies. *Iml*, the Inverse Mills ratio obtained from the multinomial logit model, is statistically different from zero both before and after SFAS 133 ($p < 0.010$ and $p < 0.002$, respectively), underscoring the necessity to control for potential self-selection bias.

We find negative and significant associations between tax-adjusted spread and both accounting measures of exposure from interest rate hedging derivatives, *IrHdg* and *DeltaIrHdg*, before and after SFAS 133.²⁴ These findings are consistent with such derivatives being used for hedging purposes and the complementary relation between derivative notional amounts and fair values in capturing derivative exposures as pointed out by Wong (2000). To test H1 and H2, we test the equality of coefficients on derivative exposure measures across pre- and post-SFAS 133 periods for the largest 25 and the remaining banks. The results of these tests are presented in Table 4, Panel B. For both subsamples, as hypothesized in H1, the associations of tax-adjusted bond spread with *IrHdg* and *DeltaIrHdg* are more negative after SFAS 133 ($p = 0.018$ and $p = 0.015$ for the largest 25 banks; $p = 0.001$ and $p = 0.032$ for the remaining banks, respectively). These findings are consistent with derivative hedges that are more effective, on average, being classified as hedging after SFAS 133.

The insignificant coefficients on the interactions of *IrHdg*, *DeltaIrHdg*, *FxHdg*, and *DeltaFxHdg* with *Big*²⁵ before and after SFAS 133 (presented in Table 4, Panel A) suggest that the pricing of hedging derivatives in bond markets does not differ between the largest 25 banks and the remaining banks.²⁵

The results in Table 4, Panel B also show that the associations between tax-adjusted spread and both *IrTrd* and *DeltaIrTrd* are more negative after SFAS 133 for the largest 25 banks ($p = 0.018$ and $p = 0.019$, respectively). Regarding the two foreign exchange derivative exposure measures (*FxTrd* and *DeltaFxTrd*), we observe a similar pattern only for *FxTrd* ($p = 0.089$). The negative coefficients on *IrTrd*, *DeltaIrTrd*, and *FxTrd* are consistent with bond markets, on average, viewing trading derivatives as being used to cross-hedge other risk exposures. The findings of greater negative coefficients on *IrTrd*, *DeltaIrTrd*, and *FxTrd* following SFAS 133 for the large banks are consistent with more economic hedges being classified as trading after SFAS 133 by these banks. These findings generally support H2. However, for the remaining banks, we find no support for H2, perhaps because these relatively smaller banks do not have much involvement in or exposures from trading derivatives. Also, except for supporting H2 for the largest 25 banks, H1 and H2 are not supported for foreign exchange rate derivatives, perhaps because banks' exposures

²³ The multinomial logit regression for *Nonuser* (*Trader*) status correctly classifies 89 percent (77 percent) of banks.

²⁴ When tax-adjusted bond spreads are decomposed into systematic risk and default risk premium components as suggested by Elton et al. (2001), we find that the negative association with spread is most pronounced for the systematic risk premium component, whereas the magnitude of the negative association with default risk premium is modest. The results of these tests are excluded for brevity.

²⁵ The negative and positive associations of tax-adjusted spread with accounting measures of derivative exposures and *Gap*, respectively, are in line with the findings of Schrand (1997) who documents that maturity gap (non-trading interest rate risk) is positively associated with higher perceived risk, and hedging derivatives decrease the maturity gap, thus reducing perceived risk.

TABLE 4
Relation between Bank Bond Spread and Derivatives before and after SFAS 133

Panel A: Association of Spread with Derivatives	Predicted Sign	Pre-SFAS 133 Period		Post-SFAS 133 Period	
		Coefficient	p-value	Coefficient	p-value
Big^{25}	-	-0.231	0.050	-0.163	0.030
$IrHdg$	-	-1.827	0.001	-2.862	0.001
$Big^{25} * IrHdg$	-/+	-0.172	0.139	-0.310	0.208
$DeltaIrHdg$	-	-10.008	0.077	-17.656	0.048
$Big^{25} * DeltaIrHdg$	-/+	-1.901	0.124	-1.366	0.242
$FxHdg$	-	-0.660	0.383	0.726	0.206
$Big^{25} * FxHdg$	-/+	0.712	0.289	0.271	0.134
$DeltaFxHdg$	-	-0.981	0.447	1.325	0.110
$Big^{25} * DeltaFxHdg$	-/+	0.235	0.268	-0.451	0.327
$IrTrd$	-/+	0.469	0.177	0.119	0.165
$Big^{25} * IrTrd$	-/+	-1.900	0.002	-2.586	0.009
$DeltaIrTrd$	-/+	-0.161	0.239	0.766	0.342
$Big^{25} * DeltaIrTrd$	-/+	-5.989	0.060	-9.354	0.048
$FxTrd$	-/+	0.409	0.207	-0.090	0.192
$Big^{25} * FxTrd$	-/+	-1.441	0.001	-1.222	0.087
$DeltaFxTrd$	-/+	-0.312	0.317	0.795	0.280
$Big^{25} * DeltaFxTrd$	-/+	-1.566	0.233	-1.169	0.392
<i>Maturity</i>	+	0.088	0.032	0.075	0.001
<i>Amount</i>	-	-0.078	0.059	-0.060	0.075
<i>Size</i>	-	-0.054	0.009	-0.078	0.001
<i>Insured</i>	+	2.885	0.154	4.061	0.143
<i>Estate</i>	+	83.889	0.288	60.009	0.204
<i>Npl</i>	+	36.419	0.005	31.593	0.001
<i>Gap</i>	+	1.998	0.009	1.834	0.001
<i>Roa</i>	-	-92.278	0.001	-71.080	0.001
<i>Cap</i>	-	-29.368	0.001	-32.721	0.006

(continued on next page)

Panel A: Association of Spread with Derivatives

	Predicted Sign	Pre-SFAS 133 Period		Post-SFAS 133 Period	
		Coefficient	p-value	Coefficient	p-value
<i>NonDerTrd</i>	+	2.800	0.007	2.599	0.013
<i>CrRisk</i>	+	10.385	0.244	3.501	0.277
<i>LoanComm</i>	-	-0.907	0.078	-1.282	0.080
<i>Guarantee</i>	-	0.091	0.260	-0.482	0.164
<i>CrDer</i>	-	-1.020	0.075	-0.866	0.161
<i>Strip</i>	-	-8.187	0.054	-5.588	0.090
<i>Iml</i>	-/+	-0.593	0.010	-0.736	0.002
Adjusted R ²			0.737		0.779
No. of Observations			270		265

Panel B: Tests of Hypotheses 1 and 2 Based on the Results Presented in Panel A

$$\begin{aligned}
 r_{it} = & \beta_0 + \beta_1 \text{Big}_t^{25} + \beta_2 \text{IrHdg}_{it} + \beta_3 \text{Big}_t^{25} * \text{IrHdg}_{it} + \beta_4 \text{DeltaIrHdg}_{it} + \beta_5 \text{Big}_t^{25} * \text{DeltaIrHdg}_{it} + \beta_6 \text{FxHdg}_{it} + \beta_7 \text{Big}_t^{25} * \text{FxHdg}_{it} + \beta_8 \text{DeltaFxHdg}_{it} \\
 & + \beta_9 \text{Big}_t^{25} * \text{DeltaFxHdg}_{it} + \beta_{10} \text{IrTrd}_{it} + \beta_{11} \text{Big}_t^{25} * \text{IrTrd}_{it} + \beta_{12} \text{DeltaIrTrd}_{it} + \beta_{13} \text{Big}_t^{25} * \text{DeltaIrTrd}_{it} + \beta_{14} \text{FxTrd}_{it} + \beta_{15} \text{Big}_t^{25} * \text{FxTrd}_{it} \\
 & + \beta_{16} \text{DeltaFxTrd}_{it} + \beta_{17} \text{Big}_t^{25} * \text{DeltaFxTrd}_{it} + \beta_{18} \text{Maturity}_{it} + \beta_{19} \text{Amount}_{it} + \beta_{20} \text{Size}_{it} + \beta_{21} \text{Insured}_{it} + \beta_{22} \text{Estate}_{it} + \beta_{23} \text{Npl}_{it} + \beta_{24} \text{Gap}_{it} \\
 & + \beta_{25} \text{NonDerTrd}_{it} + \beta_{26} \text{Cap}_{it} + \beta_{27} \text{Roa}_{it} + \beta_{28} \text{CrRisk}_{it} + \beta_{29} \text{LoanComm}_{it} + \beta_{30} \text{Guarantee}_{it} + \beta_{31} \text{CrDer}_{it} + \beta_{32} \text{Strip}_{it} + \beta_{33} \text{Iml}_{it} + \varepsilon_{it}
 \end{aligned}$$

	Largest 25 Banks		Remaining Banks	
		p-value		p-value
Hypothesis 1	$\beta_2^{\text{PRE}} + \beta_3^{\text{PRE}} > \beta_2^{\text{POST}} + \beta_3^{\text{POST}}$	0.018	$\beta_2^{\text{PRE}} > \beta_2^{\text{POST}}$	0.001
	$\beta_4^{\text{PRE}} + \beta_5^{\text{PRE}} > \beta_4^{\text{POST}} + \beta_5^{\text{POST}}$	0.015	$\beta_4^{\text{PRE}} > \beta_4^{\text{POST}}$	0.032
	$\beta_6^{\text{PRE}} + \beta_7^{\text{PRE}} > \beta_6^{\text{POST}} + \beta_7^{\text{POST}}$	0.341	$\beta_6^{\text{PRE}} > \beta_6^{\text{POST}}$	0.297
	$\beta_8^{\text{PRE}} + \beta_9^{\text{PRE}} > \beta_8^{\text{POST}} + \beta_9^{\text{POST}}$	0.205	$\beta_8^{\text{PRE}} > \beta_8^{\text{POST}}$	0.227
	$\beta_{10}^{\text{PRE}} + \beta_{11}^{\text{PRE}} > \beta_{10}^{\text{POST}} + \beta_{11}^{\text{POST}}$	0.018	$\beta_{10}^{\text{PRE}} > \beta_{10}^{\text{POST}}$	0.238
Hypothesis 2	$\beta_{12}^{\text{PRE}} + \beta_{13}^{\text{PRE}} > \beta_{12}^{\text{POST}} + \beta_{13}^{\text{POST}}$	0.019	$\beta_{12}^{\text{PRE}} > \beta_{12}^{\text{POST}}$	0.307
	$\beta_{14}^{\text{PRE}} + \beta_{15}^{\text{PRE}} > \beta_{14}^{\text{POST}} + \beta_{15}^{\text{POST}}$	0.089	$\beta_{14}^{\text{PRE}} > \beta_{14}^{\text{POST}}$	0.372
	$\beta_{16}^{\text{PRE}} + \beta_{17}^{\text{PRE}} > \beta_{16}^{\text{POST}} + \beta_{17}^{\text{POST}}$	0.392	$\beta_{16}^{\text{PRE}} > \beta_{16}^{\text{POST}}$	0.274

(continued on next page)

The sample represents observations from 1999 to 2003 (excluding 2001) for 141 banks.
All p-values are based on one-tailed t-tests when coefficient signs are predicted, and based on two-tailed tests otherwise.
All variables are as defined in Table 2 and Appendix A.

from foreign exchange rate derivatives are not material enough to affect their bond spreads.

Among the variables that proxy for exposures from off-balance sheet instruments (other than derivatives) that are potentially used for hedging, *LoanComm* and *Strip* are negatively and significantly related to tax-adjusted bond spread before and after SFAS 133 ($p = 0.078$ and $p = 0.080$ for *LoanComm*; $p = 0.054$ and $p = 0.090$ for *Strip*, respectively). *CrDer* is also negatively related to tax-adjusted spread, but only prior to SFAS 133 ($p = 0.075$).

The results reported Table 4, Panel A indicate that the association between bond spread and accounting measures of derivative exposures is driven mainly by interest rate derivatives. In unreported tests, we find support for both H1 and H2 based on different types of interest rate derivative instruments (i.e., interest rate swaps and interest rate options).

An examination of sample bank 10-Ks and Call Reports reveal that 97 of the 141 sample banks had to transfer derivatives previously classified as hedging to trading because of the various provisions of SFAS 133, indicating that SFAS 133 had direct effects on the classification of derivatives for a majority of the sample banks. Our hypotheses still hold for the banks that did not have to reclassify any derivatives. The more stringent criteria imposed by SFAS 133 (rather than the relatively lax criteria before SFAS 133) assure that derivatives classified as hedging in the post-SFAS 133 period are, in fact, highly effective hedging derivatives. In addition, as argued by the FASB (SFAS 133, ¶ 238), SFAS 133's detailed documentation requirements likely improve the monitoring of hedging activities and result in less uncertainty in the financial statements, leading to lower bank bond spreads for all derivative-user banks. Overall, these results support the view that banks benefit from the more credible derivative-related information provided under SFAS 133.

Economic Effects of SFAS 133 on Bond Spreads

The findings from the bond-pricing tests indicate that the association between the accounting measures of derivative exposures and bond spread is driven mainly by interest rate derivatives. To provide some indication of the economic effects of interest rate derivatives before and after SFAS 133, we estimate their effects on bond spread using the estimated coefficients reported in Panel A of Table 4 and descriptive statistics reported in Table 2. All estimates for the periods before and after SFAS 133 are based on mean observations for the entire sample period.

Our results indicate that banks with average values of *IrHdg* and *DeltaIrHdg* have bond spreads that are 19.5 basis points lower in the pre-SFAS 133 period and 30.8 basis points lower in the post-SFAS 133 period than banks with no exposure from interest rate hedging derivatives.²⁶ Similarly, banks with average values of *IrTrd* and *DeltaIrTrd* have bond spreads that are 5.8 basis points lower in the pre- and 9.9 basis points lower in the post-SFAS 133 period than banks with no exposure from interest rate trading derivatives.²⁷

Results of Tests of Hypotheses 3 and 4

Panels A through D of Table 5 present post-SFAS 133 and pre-SFAS 133 Spearman correlations (above and below the diagonal, respectively) for the variables used in the exposure co-variation tests. First, we partition the entire sample into banks with large and small portfolios of hedging derivatives.²⁸ Panels A and B display correlations among the variables used to estimate

²⁶ We use the weighted average of the coefficients for the largest 25 banks and other banks when estimating the effects of hedging derivatives because the effect of hedging derivatives on bond spread is significantly negative for both groups of banks.

²⁷ We estimate the effects of trading derivatives only for the largest 25 banks because our results indicate that the effect of trading derivatives on bond spread is significantly negative for only these banks.

²⁸ We consider a bank to have a large (small) portfolio of hedging derivatives if the notional amount of hedging derivatives scaled by beginning book value of equity is above (below) the sample median in the pre-SFAS 133 period.

TABLE 5
Association of Derivative and Non-Derivative Exposures before and after SFAS 133
Correlation Analyses

Panel A: Correlations for High Users of Hedging Derivatives			
	<u>$\Delta DerTrading$</u>	<u>$\Delta DerHedging$</u>	<u>$\Delta NonDerTrading$</u>
$\Delta DerTrading$		-0.049 (0.245)	-0.103 (0.015)
$\Delta DerHedging$	-0.019 (0.655)		-0.051 (0.229)
$\Delta NonDerTrading$	-0.062 (0.140)	0.033 (0.431)	
$\Delta NonTrading$	-0.057 (0.177)	-0.227 (0.001)	0.046 (0.272)
Panel B: Correlations for Low Users of Hedging Derivatives			
	<u>$\Delta DerTrading$</u>	<u>$\Delta DerHedging$</u>	<u>$\Delta NonDerTrading$</u>
$\Delta DerTrading$		0.069 (0.102)	-0.068 (0.106)
$\Delta DerHedging$	0.005 (0.915)		0.033 (0.441)
$\Delta NonDerTrading$	-0.044 (0.297)	-0.009 (0.830)	
$\Delta NonTrading$	0.027 (0.528)	-0.155 (0.001)	0.030 (0.474)
Panel C: Correlations for High Users of Trading Derivatives			
	<u>$\Delta DerTrading$</u>	<u>$\Delta DerHedging$</u>	<u>$\Delta NonDerTrading$</u>
$\Delta DerTrading$		-0.019 (0.758)	-0.128 (0.039)
$\Delta DerHedging$	-0.031 (0.657)		-0.034 (0.587)
			0.044 (0.479)
			-0.251 (0.001)

(continued on next page)

Panel C: Correlations for High Users of Trading Derivatives

	$\Delta DerTrading$	$\Delta DerHedging$	$\Delta NonDerTrading$	$\Delta NonTrading$
$\Delta NonDerTrading$	-0.091 (0.196)	0.089 (0.207)		0.050 (0.418)
$\Delta NonTrading$	0.021 (0.772)	-0.204 (0.004)	-0.007 (0.925)	

Panel D: Correlations for Low Users of Trading Derivatives

	$\Delta DerTrading$	$\Delta DerHedging$	$\Delta NonDerTrading$	$\Delta NonTrading$
$\Delta DerTrading$		0.090 (0.147)	-0.073 (0.238)	-0.022 (0.718)
$\Delta DerHedging$	0.045 (0.526)		0.035 (0.570)	-0.190 (0.002)
$\Delta NonDerTrading$	-0.056 (0.426)	-0.061 (0.389)		0.040 (0.522)
$\Delta NonTrading$	0.078 (0.273)	-0.148 (0.035)	-0.005 (0.940)	

Correlation coefficients in Panels A and B are based on financial data from 1999 to 2003 (excluding 2001) for 141 banks.

Correlation coefficients in Panels C and D are based on financial data for 52 banks in the pre-SFAS 133 and 67 banks in the post-SFAS 133 period.

Post-SFAS 133 and pre-SFAS 133 Spearman correlations are presented above and below the diagonal, respectively.

Significance levels in parentheses are based on two-tailed tests.

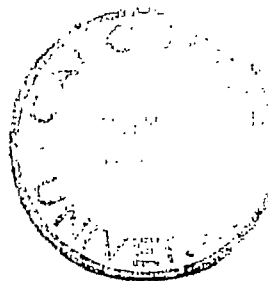
Variable Definitions:

$\Delta DerTrading$ = quarterly change in the fair value of net trading derivative assets scaled by beginning book value of equity;

$\Delta DerHedging$ = quarterly change in the fair value of net hedging derivative assets scaled by beginning book value of equity;

$\Delta NonDerTrading$ = quarterly change in the fair value of net non-derivative trading assets scaled by beginning book value of equity; and

$\Delta NonTrading$ = quarterly change in the fair value of net non-trading assets scaled by beginning book value of equity.



Equation (3). As presented in Panel A, for high-hedger banks, the correlation between $\Delta DerHedging$ and $\Delta NonTrading$ is more negative after SFAS 133 (i.e., it changes from -0.227 to -0.284), providing preliminary evidence in support of H3. Although not as sizable, the change in the correlation coefficient (from -0.155 to -0.175) is in the same direction for banks with small portfolios of hedging derivatives as presented in Panel B.

Panels C and D of Table 5 display correlations among the variables used in estimating Equation (4) for banks with large and small portfolios of trading derivatives, respectively. The negative correlations between $\Delta DerTrading$ and $\Delta NonDerTrading$ presented in Panels C and D provide evidence in support of trading derivatives offsetting exposures from non-derivative trading positions. For high-trader banks, which primarily consist of large banks, the correlation between $\Delta DerTrading$ and $\Delta NonDerTrading$ is more negative after the adoption of SFAS 133 ($r = -0.128$; $p = 0.039$) compared to the pre-SFAS 133 period ($r = -0.091$; $p = 0.196$), consistent with more economic hedges being classified as trading after SFAS 133. A similar trend also exists for low-trader banks. However, the correlation coefficients are smaller and insignificant ($r = -0.056$; $p = 0.426$ and $r = -0.073$; $p = 0.238$ before and after SFAS 133, respectively).

The insignificant correlation coefficients between $\Delta DerHedging$ and $\Delta NonDerTrading$ ($\Delta DerTrading$ and $\Delta NonTrading$) observed in Panels A through D indicate that banks do not cross-hedge trading (non-trading) positions with hedging (trading) derivatives. These findings are consistent with Liu et al. (2004), who find no significant cross-hedging between trading exposures and non-trading exposures.

The results from estimating Equation (3) that are reported in Panel A of Table 6 are mostly consistent with the correlation analyses. Panel C presents the results of the tests of H3 for banks with large and small portfolios of hedging derivatives. The coefficient on $\Delta NonTrading$ is more negative after SFAS 133 for both groups of banks ($p = 0.030$ and $p = 0.078$, respectively), providing further evidence that derivatives classified as hedging consist of hedges that meet stricter hedge effectiveness criteria after SFAS 133.

Results from estimating Equation (4) are reported in Panel B of Table 6. Panel C presents the results of the tests of H4 for banks with large and small portfolios of trading derivatives. The negative and significant coefficient on $\Delta NonDerTrading$ for both groups of banks before and after SFAS 133 suggests that trading derivatives cross-hedge non-derivative trading exposures. Also, the coefficient on $\Delta NonDerTrading$ is more negative subsequent to SFAS 133 for banks with large portfolios of trading derivatives ($p = 0.054$), providing support for H4. More specifically, this finding suggests that the extent of cross-hedging between trading derivatives and non-derivative trading positions is greater after the adoption of SFAS 133, consistent with more economic hedges that do not qualify as accounting hedges being classified as trading subsequent to SFAS 133. However, for banks with small portfolios of trading derivatives, the coefficients on $\Delta NonDerTrading$ are not reliably different before and after SFAS 133 ($p = 0.151$).

Sensitivity Analyses and Additional Considerations

We conduct several sensitivity analyses to assess the robustness of our findings. First, we re-run our tests after deleting the three largest banks (J.P. Morgan Chase, Bank of America, and Citibank) that dominate the derivatives market during our sample period. We continue to find support for all hypotheses. Second, a potential explanation for our findings is that the change in the yield curve coincides with the adoption of SFAS 133. The flat yield curve during the pre-SFAS 133 period (1999–2000) became steeply upward sloped in 2001. It remained so through the end of 2004 and flattened in the 2005–2006 period. In order to address the possibility that our results are driven by the different-shaped yield curves in the pre- and post-SFAS 133 periods rather than by SFAS 133, we re-test our hypotheses using the years 2005–2006 as the post-SFAS 133 period. This specification ensures that both pre- and post-SFAS 133 period yield curves are flat. We

TABLE 6
Association of Derivative and Non-Derivative Exposures before and after SFAS 133
Regression Analyses

Panel A: Hedging Derivatives and Non-Derivative Exposures

	Pre-SFAS 133 Period		Post-SFAS 133 Period	
	Coefficient	p-value	Coefficient	p-value
$Hight^{Hds}$	0.010	0.183	0.009	0.253
$\Delta NonDerTrading$	0.010	0.462	-0.007	0.358
$High^{Hds} * \Delta NonDerTrading$	0.021	0.657	0.003	0.405
$\Delta NonTrading$	-0.108	0.022	-0.144	0.006
$Hight^{Hds} * \Delta NonTrading$	-0.051	0.005	-0.081	0.041
Adjusted R ²		0.182		0.218
No. of Observations		1,048		1,014

Panel B: Trading Derivatives and Non-Derivative Exposures

	Pre-SFAS 133 Period		Post-SFAS 133 Period	
	Coefficient	p-value	Coefficient	p-value
$Hight^{Trd}$	-0.003	0.284	0.012	0.176
$\Delta NonDerTrading$	-0.056	0.038	-0.070	0.008
$Hight^{Trd} * \Delta NonDerTrading$	-0.030	0.021	-0.071	0.064
$\Delta NonTrading$	-0.034	0.412	0.093	0.222
$Hight^{Trd} * \Delta NonTrading$	0.041	0.255	0.002	0.334
Adjusted R ²		0.101		0.134
No. of Observations		404		510

(continued on next page)

Panel C: Tests of Hypotheses 3 and 4 Based on the Results Presented in Panels A and B

Hypothesis	High Users of Hedging Derivatives	p-value	Low Users of Hedging Derivatives	p-value
	$\gamma_4^{PRE} + \gamma_5^{PRE} > \gamma_4^{POST} + \gamma_5^{POST}$		$\gamma_4^{PRE} > \gamma_4^{POST}$	
Hypothesis 3		0.030		0.078
Hypothesis	High Users of Trading Derivatives	p-value	Low Users of Trading Derivatives	p-value
	$\lambda_2^{PRE} + \lambda_3^{PRE} > \lambda_2^{POST} + \lambda_3^{POST}$		$\lambda_2^{PRE} > \lambda_2^{POST}$	
Hypothesis 4		0.054		0.151

All results are based on standard errors clustered simultaneously by bank and by year.

All p-values are based on two-tailed t-tests.

All other variables are as defined in Table 5.

Variable Definition:

$High^{Hdg}$

$(High^{Hdg}) = 1$ if the notional amount of hedging (trading) derivatives scaled by beginning book value of equity is above sample median for the year, and 0 otherwise.

continue to find strong support for our hypotheses, suggesting that the change in the yield curve in 2001 is not driving our results. Third, as mentioned earlier, we estimate the size of the net long/short positions in trading derivatives for some banks. These estimates are likely to contain error. Therefore, we re-run our tests using gross (long plus short) notional values for trading derivatives. Our results regarding the effect of SFAS 133 still hold. Fourth, we repeat our tests after excluding the top and bottom 1 percent of the distribution of each variable to assess the sensitivity of our findings to extreme observations. We continue to observe results consistent with our primary findings. Last, we find similar results for the exposure covariation tests of H3 and H4 when we scale all variables by beginning total assets instead of beginning book value of equity as we do in our primary tests.

Another potential alternative explanation for the difference between the risk relevance of derivatives before and after SFAS 133 is the possibility of a change in banks' derivative strategies. However, from a cost-benefit standpoint, we believe that there are a number of reasons for banks not to change their derivative strategies following SFAS 133. First, derivatives are efficient and low-cost risk management tools for banks. In the absence of derivatives, managing interest rate risk would require significant amounts of capital investment. In addition, derivative contracts can be customized, depending on the risk being hedged. Thus, the benefits of managing interest rate risk via efficient and low-cost risk management tools likely outweigh the costs of using alternative hedging strategies. Second, a major concern about SFAS 133 expressed by banks is the increased volatility likely to be introduced by SFAS 133 to banks' earnings.²⁹ However, banks can potentially offset this volatility in a cost-effective way by using other tools for smoothing income (such as loan loss provisions) without changing their derivative strategies (Wahlen 1994; Liu and Ryan 2006; Kanagaretnam et al. 2010). In fact, Kilic et al. (2010) document that banks that are likely to be most impacted by SFAS 133 have increased their reliance on loan loss provisions for income smoothing in the post-SFAS 133 period.

Finally, we note that, although the volume of trading derivatives grew rapidly during our sample period,³⁰ this rapid growth is unlikely to drive our results regarding the effects of SFAS 133 for a number of reasons. First, our hypotheses are stated and models designed to test whether the pricing of a unit of derivative exposure changed with SFAS 133, regardless of the amount of total exposure. Second, the growth trend cannot be generalized to all banks; it is driven mostly by the largest seven derivative-dealer banks (Office of the Comptroller of the Currency Report on Derivatives Activities 1995–2006). We continue to find support for all our hypotheses when we re-run our tests after deleting these seven banks from the sample. Third, the growth trend started in the early 1990s and has continued since then. Thus, there is no evidence of a shift in the trend that coincides with the adoption of SFAS 133.

V. SUMMARY AND CONCLUSION

SFAS 133 is a widely debated and controversial accounting standard that replaced the incomplete and inconsistent accounting guidance for derivatives that existed prior to its adoption. We examine how SFAS 133 affected the risk relevance of accounting measures of bank derivative exposures to bond investors. We hypothesize and find that interest rate hedging derivatives are more negatively associated with bond spreads after SFAS 133. These findings are consistent with

²⁹ SFAS 133 was expected to introduce more volatility to banks' earnings because it requires more economic hedges to be classified as trading with gains/losses recognized in income immediately and any hedge ineffectiveness to be recognized in income immediately.

³⁰ The annual growth rates of derivatives classified as trading are 35.71 percent in 1996, 29.46 percent in 1997, 53.65 percent in 1998, 3.25 percent in 1999, 17.88 percent in 2000, 11.92 percent in 2001 (the year of adoption of SFAS 133), 21.32 percent in 2002, 28.76 percent in 2003, 28.61 percent in 2004, 5.38 percent in 2005, and 11.52 percent in 2006.

SFAS 133 restricting derivatives classified as hedging to only those instruments that serve as highly effective and documented derivative hedges. For the largest 25 banks, which are extensively involved in derivatives trading, we find that interest and foreign exchange rate trading derivatives are more *negatively* associated with bond spreads after SFAS 133, consistent with more economic hedges being classified as trading after SFAS 133. To provide further insights, we also examine how SFAS 133 altered the association of exposures from derivatives with non-derivative exposures. Consistent with our results from the risk-relevance tests, the findings from the exposure covariation tests indicate that the offsetting effects between hedging derivatives and non-trading exposures, and between trading derivatives and non-derivative trading exposures, are greater after SFAS 133. Taken together, these findings indicate that, despite its limitations, SFAS 133 has improved the transparency and the monitoring of the risk implications of derivatives.

APPENDIX A CONTROL VARIABLES USED IN THE BOND-PRICING MODELS

Variable Name	Data Source	Definition and Explanation
Bond Characteristics		
<i>Maturity</i>	Mergent FISD	Natural logarithm of the bond's years to maturity. Bonds with longer maturities (<i>MATURITY</i>) are expected to have higher spreads due to lower liquidity (Hancock and Kwast 2001; Covitz et al. 2000).
<i>Amount</i>	Mergent FISD	Natural logarithm of the size of the bond issue in millions of dollars. Bonds issued in larger amounts are expected to have lower spreads due to higher liquidity benefits (Hancock and Kwast 2001; Morgan and Stiroh 2001) and due to economies of scale in underwriting (Sengupta 1998).
Bank Characteristics		
<i>Size</i>	FR 9-YC	Natural logarithm of the bank's total assets. Larger banks are less likely to default because they are better diversified and are more likely to benefit from the "too-big-to-fail" policy formalized by the Comptroller of the Currency (Flannery and Sorescu 1996; Morgan and Stiroh 2001). Also, their bonds are more liquid and, therefore, are traded at lower spreads over Treasury bonds (Flannery and Sorescu 1996).
<i>Insured</i>	Bank Compustat	Ratio of total insured deposits to total deposits. Banks with larger insured deposits are subject to less market monitoring and discipline (Billett et al. 1998) and, therefore, shift to greater use of insured deposits as their financial condition deteriorates (Billett et al. 1998; Jordan 2000). As a result, banks with higher insured deposits are perceived to have greater potential for moral hazard and, therefore, higher bond spreads (Jagtiani et al. 2002).
<i>Estate</i>	FR 9-YC	Real estate acquired in satisfaction of debt previously contracted scaled by total assets. Creditors may view higher property collected on foreclosed real estate loans as a signal of lower loan quality and higher credit risk, and may demand higher bond spread (Morgan and Stiroh 2001).

(continued on next page)

Variable Name	Data Source	Definition and Explanation
<i>Npl</i>	FR 9-YC	Sum of nonaccrual loans and accruing loans past due 90 days or more scaled by total assets. Higher nonaccrual and past due loans signal lower loan quality and higher loan losses. Higher loan losses reduce the capital base more and require greater charges against income through higher provisions for loan losses, thereby increasing bank default risk (Flannery and Sorescu 1996; Morgan and Stiroh 2001).
<i>Gap</i>	10-K	Absolute value of the difference between bank non-trading assets and non-trading liabilities maturing or repricing within one year scaled by total assets. Higher maturity mismatch between non-trading assets and liabilities increases a bank's non-trading interest rate risk exposure and overall risk (Flannery and James 1984; Flannery and Sorescu 1996; Morgan and Stiroh 2001).
<i>NonDerTrd</i>	10-K	Fair value of non-derivative trading assets scaled by total assets. Trading activities are associated with higher bond spreads due to potential agency risks associated with such activities. That is, banks may more easily assume trading risks at levels that are suboptimal for bondholders, since trading positions can be easily changed or concealed with leverage. Such uncertainty about trading activities gives rise to agency risks (Morgan and Stiroh 2001).
<i>Cap</i>	FR Y9-C	Ratio of book value of equity to risk-weighted assets. A bank's assets differ in terms of riskiness. Some assets (e.g., commercial loans) are considered riskier than others (e.g., real estate loans), and are given higher risk weights in the calculation of regulatory capital. Bank capital is a cushion for potential losses. Thus, <i>Cap</i> will be lower for banks with riskier assets. <i>Cap</i> jointly captures the asset risk composition and capital adequacy. Higher capital is likely to result in creditors demanding lower spread.
<i>Roa</i>	FR Y9-C	Annual net income scaled by total assets. Higher profitability is associated with lower likelihood of default and bond spread (Flannery and Sorescu 1996; Morgan and Stiroh 2001).
<i>CrRisk</i>	FR Y9-C	Charge-offs incurred on derivative contracts during the year (i.e., losses incurred when the counterparty fails to meet the settlement obligations of the derivative contract) scaled by total assets. Our derivative exposure measures proxy market risk exposure. However, derivatives also expose the bank to credit risk. The charge-offs incurred on derivatives are expected to increase bond spread.
Off-Balance Sheet Hedges other than Market Risk Derivatives		
<i>LoanComm</i>	10-K	Amount of loans for which official promises to lend are conveyed scaled by total assets. Banks sometimes use forward loan sales commitments as economic hedges of commitments to originate loans (Office of the Comptroller of the Currency 2005) because fair values of these commitments are offsetting.

(continued on next page)

Variable Name	Data Source	Definition and Explanation
<i>Guarantee</i>	10-K	On- and off-balance sheet guarantees and insurance contracts that back financial obligations scaled by total assets. Banks enter into financial guarantee and insurance contracts that back off-balance sheet instruments or financial obligations such as outstanding securities and loans.
<i>CrDer</i>	10-K	Notional amount of credit derivatives on which the bank is the beneficiary scaled by total assets. In order to hedge their credit risks (losses) from the loans they originate, banks sometimes enter into credit derivative contracts with the bank as the beneficiary. Such contracts constitute guarantees against a bank's credit losses from other parties and are expected to reduce bank credit risk and thus bond spread.
<i>Strip</i>	10-K	Fair value of principal-only strips scaled by total assets. Principal-only strips do not qualify as derivative instruments under the requirements of SFAS 133 (paragraph 14), but these instruments are used to hedge prepayment risk especially from mortgage servicing rights, which is driven by decreasing interest rates.

APPENDIX B

VARIABLES USED IN THE FIRST-STAGE MULTINOMIAL LOGIT MODEL

Variable Name	Data Source	Definition and Explanation
Big_{it}^{25}	FR Y-9C	Dummy variable that equals 1 if the bank is one of largest 25 banks in the United States, and 0 otherwise. Holdings of derivatives in the United States are concentrated in the largest 25 banks, whereas the remaining holdings (particularly for hedging purposes) are still economically significant for the remaining user banks because their asset portfolios are remarkably smaller than those of the largest 25 banks. In order to better capture the effect of large banks, we include discrete size variables.
$Big_{it}^3(Big_{it}^{4-25})$	FR Y-9C	Dummy variable that equals 1 if the bank is one of largest three (4 to 25) banks in the United States, and 0 otherwise. Among the largest 25 banks, the largest three banks (J.P. Morgan Chase, Bank of America, and Citigroup) dominate derivatives activity (Office of the Comptroller of the Currency Derivatives Report, 1998–2003).
<i>Size</i>	FR Y-9C	Natural logarithm of total assets. A number of studies (e.g., Booth et al. 1984; Koppenhaver 1990; Kim and Koppenhaver 1992; Gunther and Siems 2002) document that bank size is positively related to derivative use because large banks can more easily bear costs associated with implementing derivative trading and hedging programs such as the costs of personnel training and internal control systems.

(continued on next page)

Variable Name	Data Source	Definition and Explanation
<i>Lev</i>	FR Y-9C	Ratio of risk-weighted assets to book value of equity. <i>Lev</i> jointly captures the probability of financial distress (Nance et al. 1993; Géczy et al. 1997; Sinkey and Carter 2000) and balance sheet composition in terms of asset riskiness. Banks with higher leverage have higher probability of financial distress, and are more likely to use derivatives to reduce the variance of bank value. Also, banks holding assets with higher credit risks (i.e., higher risk-weighted assets and therefore higher <i>Lev</i>) are more likely to use derivatives to hedge other exposures (Schrand and Unal 1998). On the other hand, a bank with low capital may lack credibility as counterparty and creates credit risk for the opposite party in derivatives trading. Therefore, the likelihood of holding trading derivatives is expected to decrease with <i>Lev</i> (Kim and Koppenhaver 1992).
<i>Gap</i>	FR Y-9C	Absolute value of the difference between bank assets and liabilities maturing or repricing within one year scaled by total assets. Banks with greater maturity mismatch and, therefore, greater exposure to interest rate risk are more likely to use derivatives to hedge their exposures (Sinkey and Carter 2000; Gunther and Siems 2002).
<i>Nim</i>	FR Y-9C	Net interest income scaled by total assets. Banks with lower net interest margins are more likely to use derivatives, particularly for trading purposes, in an attempt to increase fee income by selling derivatives or by speculating in derivative markets to offset reduced spreads from traditional banking activities (Shyu and Reichert 2002).
<i>Npl</i>	FR Y-9C	Ratio of the sum of nonaccrual loans and accruing loans past due 90 days or more to total assets. Lower loan quality is associated with higher credit risk. As part of a joint-risk management program, banks exposed to significant levels of credit risk are more likely to use derivatives to hedge other exposures (Schrand and Unal 1998).
<i>Liquidity</i>	FR Y-9C	Sum of cash and due, federal funds sold and securities purchased to resell scaled by total assets. Banks with greater liquidity are in a better position to manage large derivative positions (Jordan 1995) and are more likely to engage in derivatives trading activities.
<i>Incentive</i>	ExecuComp	Ratio of option vega to delta. Risk management theory identifies risk-taking incentives of managers as a determinant of risk management (Stulz 1984; Smith and Stulz 1985). Following Core and Guay (2002), we measure manager incentives to increase risk (stock price) by vega (delta), and include <i>Incentive</i> to capture the effect of managerial risk-taking incentives on the decision to use derivatives (Tufano 1996; Rogers 2002).

(continued on next page)

Variable Name	Data Source	Definition and Explanation
<i>Convexity</i>	Bank Compustat	Expected percentage tax savings from a 5 percent reduction in the volatility of taxable income. In the presence of tax function convexity, income smoothing reduces expected tax liabilities (Smith and Stulz 1985). Hedging serves as a mechanism to smooth income over time. Therefore, a number of studies (e.g., Nance et al. 1993; Dolde 1995; Mian 1996; Tufano 1996; Géczy et al. 1997; Graham and Rogers 2002) control for tax convexity as a determinant of derivative usage for hedging purposes. We measure tax convexity following Graham and Smith (1999).
<i>ComIndLoan</i>	FR Y-9C	Ratio of commercial and industrial loans to total assets. Commercial and industrial loan customers use derivatives to hedge their risks, and banks' familiarity with potential derivative customers increase the likelihood and extent of their participation in derivatives market as a dealer (Kim and Koppenhaver 1992).

APPENDIX C DERIVATIONS OF UNADJUSTED SPREAD AND TAX-ADJUSTED SPREAD

Unadjusted Spread

We compute unadjusted bond yield spreads as the difference between the actual yield to maturity on the bank bond at issuance, and the yield to maturity on a synthetic Treasury bond, r_s , that perfectly matches the bank bond in terms of coupon payment, maturity, duration, and convexity.

In order to obtain r_s , we first treat each payment made by a coupon-bearing bank bond as a zero-coupon bond. For each of these payments, we calculate the risk-free spot rate using a parametric yield curve specification suggested by Gurkaynak et al. (2006). This specification, initially proposed by Nelson and Siegel (1987) and extended by Gurkaynak et al. (2006), is as follows:

$$r_t(n) = \delta_0 + \delta_1 \frac{1 - e\left(\frac{-n}{\tau_1}\right)}{\frac{n}{\tau_1}} + \delta_2 \left[\frac{1 - e\left(\frac{-n}{\tau_1}\right)}{\frac{n}{\tau_1}} - e\left(\frac{-n}{\tau_1}\right) \right] + \delta_3 \left[\frac{1 - e\left(\frac{-n}{\tau_2}\right)}{\frac{n}{\tau_2}} - e\left(\frac{-n}{\tau_2}\right) \right], \quad (\text{A.1})$$

where $r_t(n)$ is the yield to maturity of an n -year zero-coupon risk-free (Treasury) bond at time t ; n is number of years till maturity; τ_1 and τ_2 are parameters that determine the locations of the "humps" in the yield curve. The parameters of this specification from 1961 to 2006 are provided on a daily basis by Gurkaynak et al. (2006).³¹ Thus, for a given date during the 1961–2006 period, yields can be computed for future bond payments, each of which is assumed to be a zero-coupon bond, at all maturities. Next, we discount each payment made by the bond at a rate derived from Equation (A.1), and sum up the discounted values to obtain the price of a Treasury bond, which perfectly matches the bank bond in terms of coupon payment, maturity, duration, and convexity. The Treasury bond price, P^{TR} , is computed as:

³¹ The parameters are also available at <http://www.federalreserve.gov/pubs/feds/2006>.

$$P^{TR} = \frac{C}{(1+r_1)^1} + \frac{C}{(1+r_2)^2} + \dots + \frac{C}{(1+r_n)^n} + \frac{F}{(1+r_n)^n}, \quad (\text{A.2})$$

where r_n is the yield derived from Equation (A.1) for a zero-coupon bond that matures in n years; C is the coupon payment; and F is the face value of the bond. In order to obtain the appropriate risk-free spot rate, we then solve the following equation for r_s :

$$P^{TR} = \frac{C}{(1+r_s)^1} + \frac{C}{(1+r_s)^2} + \dots + \frac{C}{(1+r_s)^n} + \frac{F}{(1+r_s)^n}. \quad (\text{A.3})$$

Tax-Adjusted Spread (Spread less Tax Premium)

In order to obtain a measure of tax-adjusted spread, r_{slt} , we reduce total spread, r_s , by the amount of estimated tax premium, r_t . We estimate r_t as the difference between the portion of spread explained by expected losses from default and taxes, and the portion explained by default premium only. Re-arranging the equations in Elton et al. (2001) yields:

$$r_t = \ln \left[\frac{D}{D-T} \right], \quad (\text{A.4})$$

where $D = (1 - P_{t+1}) + \frac{\phi P_{t+1}}{V_{t+1}T+C}$, and $T = \frac{[C(1-P_{t+1}) - (1-\phi)P_{t+1}]}{V_{t+1}T+C} t_s(1-t_f)$, and r_t is the tax premium; t_s is the state tax rate; t_f is the federal tax rate; and other terms are as defined in Equations (A.2) and (A.3). Spread less tax premium, r_{slt} , then equals the difference between r_s and r_t . We assume an effective tax rate, $t_s(1-t_f)$, of 4.2 percent based on the grid search procedure in Elton et al. (2001). We examine different levels of tax rates between 3 percent and 5 percent at increments of 0.2 percent.³² An effective tax rate of 4.2 percent minimizes the squared mean error between derived and actual sample bond prices. This effective tax rate is consistent with pairs of federal and state tax rates such as 33 percent and 6.3 percent, respectively.

REFERENCES

- Ahmed, A. S., E. Kilic, and G. Lobo. 2006. Does recognition versus disclosure matter? Evidence from value-relevance of banks' recognized and disclosed derivative financial instruments. *The Accounting Review* 81 (3): 567-588.
- America's Community Bankers. 1997. *Comment Letter to FASB Regarding SFAS 133*. Letter No. 132. Norwalk, CT: FASB.
- Anson, M. J. P. 1999. *Accounting and Tax Rules for Derivatives*. New Hope, PA: Frank J. Fabozzi Associates.
- Bank of America. 2006. *A Survey of the Effects of SFAS 133 on Corporate Hedging Practice*. Charlotte, NC: Bank of America.
- Barth, M. E., W. H. Beaver, and W. R. Landsman. 1996. Value-relevance of banks' fair value disclosures under SFAS No. 107. *The Accounting Review* 71 (4): 513-537.
- Beier, R. 1993. Will the FASB clip the hedges? *Financial Executive* 9 (4): 40-43.
- Beaver, W., P. Kettler, and M. Scholes. 1970. The association between market-determined and accounting-determined measures of risk. *The Accounting Review* 45 (4): 654-682.
- Bhojraj, S., and B. Swaminathan. 2009. How does the corporate bond market value capital investments and accruals? *Review of Accounting Studies* 14 (1): 31-62.
- Billett, M. T., J. A. Garfinkel, and E. S. O'Neal. 1998. The cost of market versus regulatory discipline in banking. *Journal of Financial Economics* 48 (3): 333-358.

³² This range is determined based on Elton et al.'s (2001) finding of an effective tax rate of 4 percent for their sample.

- Booth, J. R., R. L. Smith, and R. W. Stolz. 1984. Use of interest rate futures by financial institutions. *Journal of Bank Research* 15 (1): 15–20.
- Canadian Bankers Association. 1997. *Comment Letter to FASB Regarding SFAS 133*. Letter No. 36. Norwalk, CT: FASB.
- Carpenter, J. N. 1996. Current issues in accounting for derivatives. *Journal of Derivatives* 3 (3): 65–71.
- Consumer Federation of America. 1997. *Comment Letter to FASB Regarding SFAS 133*. Letter No. 18. Norwalk, CT: FASB.
- Core, J., and W. Guay. 2002. Estimating the value of employee stock option portfolios and their sensitivities to price and volatility. *Journal of Accounting Research* 40 (3): 613–630.
- Covitz, D. M., D. Hancock, and M. L. Kwast. 2000. *Mandatory Subordinated Debt: Would Banks Face More Market Discipline?* Washington, D.C.: Board of Governors of the Federal Reserve System.
- Dadalt, P., G. Gay, and J. Nam. 2002. Asymmetric information and corporate derivatives use. *Journal of Futures Markets* 22 (3): 241–267.
- Dolde, W. 1995. Hedging, leverage and primitive risk. *Journal of Financial Engineering* 4 (2): 187–216.
- Eccher, A., K. Ramesh, and S. R. Thiagarajan. 1996. Fair value disclosures of bank holding companies. *Journal of Accounting and Economics* 22 (1–3): 79–117.
- Elton, E. J., M. J. Gruber, D. Agrawal, and C. Mann. 2001. Explaining the rate spread on corporate bonds. *The Journal of Finance* 56 (1): 247–277.
- Financial Accounting Standards Board (FASB). 1981. *Foreign Currency Translation*. Statement of Financial Accounting Standards No. 52. Stamford, CT: FASB.
- . 1984. *Accounting for Futures Contracts*. Statement of Financial Accounting Standards No. 80. Stamford, CT: FASB.
- . 1994. *Disclosure about Derivative Financial Instruments and Fair Value of Financial Instruments*. Statement of Financial Accounting Standards No. 119. Stamford, CT: FASB.
- . 1998. *Accounting for Derivative Instruments and Hedging Activities*. Statement of Financial Accounting Standards No. 133. Stamford, CT: FASB.
- . 2000. *Accounting for Certain Derivative Instruments and Certain Hedging Activities—An Amendment of FASB Statement No. 133*. Statement of Financial Accounting Standards No. 138. Stamford, CT: FASB.
- . 2003. *Amendment of Statement No. 133 on Derivative Instruments and Hedging Activities*. Statement of Financial Accounting Standards No. 149. Stamford, CT: FASB.
- Financial Executives Institute. 1997. *Comment Letter to FASB Regarding SFAS 133*. Letter No. 21. Norwalk, CT: FASB.
- Finnerty, J. D., and D. Grant. 2002. Alternative approaches to testing hedge effectiveness under SFAS No. 133. *Accounting Horizons* 16 (2): 95–108.
- First Chicago NBD Corporation. 1997. *Comment Letter to FASB Regarding SFAS 133*. Letter No. 104. Norwalk, CT: FASB.
- Flannery, M. J., and C. M. James. 1984. The effect of interest rate changes on the common stock returns of financial institutions. *The Journal of Finance* 39 (4): 1141–1153.
- , and S. Sorescu. 1996. Evidence of bank market discipline in subordinated debenture yields: 1983–1991. *The Journal of Finance* 51 (4): 1347–1377.
- Francis, J., C. Lennox, and Z. Wang. 2010. Selection models in accounting research. Working paper, University of Missouri–Columbia and Nanyang Technological University.
- Froot, K. A., D. S. Scharfstein, and J. C. Stein. 1993. Risk management: Coordinating corporate investment and financing policies. *The Journal of Finance* 48 (5): 1629–1658.
- Gastineau, G., D. Smith, and R. Todd. 2001. *Risk Management, Derivatives, and Financial Analysis under SFAS No. 133*. Charlottesville, VA: Blackwell Publishers.
- Gay, G. D., and J. Nam. 1998. The underinvestment problem and corporate derivatives use. *Financial Management* 27 (4): 53–69.
- Géczy, C., B. A. Minton, and C. Schrand. 1997. Why firms use currency derivatives. *The Journal of Finance* 52 (4): 1323–1354.
- Graham, J. R., and D. A. Rogers. 2002. Do firms hedge in response to tax incentives? *The Journal of Finance* 57 (2): 815–839.

- Graham, J. R., and C. W. Smith. 1999. Tax incentives to hedge. *The Journal of Finance* 54 (6): 2241–2262.
- Gunther, J. W., and T. F. Siems. 2002. The likelihood and extent of banks' involvement with interest rate derivatives as end users. *Research in Finance* 19 (1): 125–142.
- Gurkaynak, R. S., B. Sack, and J. H. Wright. 2006. *The U.S. Treasury Yield Curve: 1961 to the Present*. Finance and Economics Discussion Series, Divisions of Research and Statistics and Monetary Affairs. Washington, D.C.: Federal Reserve Board.
- Hancock, D., and M. L. Kwast. 2001. Using subordinated debt to monitor bank holding companies: Is it feasible? *Journal of Financial Services Research* 20 (2–3): 147–187.
- Hodder, L., P. Hopkins, and J. Wahlen. 2006. Risk-relevance of fair-value income measures for commercial banks. *The Accounting Review* 81 (2): 337–375.
- Hull, J. 1993. *Options, Futures, and Other Derivative Securities*. Englewood Cliffs, NJ: Prentice Hall.
- Jagtiani, J., G. Kaufman, and C. Lemieux. 2002. The effect of credit risk on bank and bank holding company bond yields: Evidence from the post-FDICIA period. *Journal of Financial Research* 25 (4): 559–575.
- Jordan, J. 1995. Supervision of derivative instruments. *Journal of Financial Services Research* 9 (3–4): 433–444.
- . 2000. *Depositor Discipline at Failing Banks*. Boston, MA: Federal Reserve Bank of Boston.
- Kanagaretnam, K., G. V. Krishnan, and G. J. Lobo. 2010. An empirical analysis of auditor independence in the banking industry. *The Accounting Review* 85 (6): 1–36.
- Kilic, E., G. J. Lobo, T. Ranasinghe, and K. Sivaramakrishnan. (2010). Earnings management and value relevance consequences of SFAS 133: Evidence from bank holding companies. Working paper, University of Houston, Houston, TX.
- Kim, S., and G. D. Koppenhaver. 1992. An empirical analysis of bank interest rate swaps. *Journal of Financial Services Research* 7 (1): 57–72.
- Koppenhaver, G. D. 1990. An empirical analysis of bank hedging in futures markets. *Journal of Futures Markets* 10 (1): 1–12.
- Lee, L. 1983. Generalized econometric models with selectivity. *Econometrica* 51 (2): 507–512.
- Liu, C., S. Ryan, and H. Tan. 2004. How banks' value-at-risk disclosures predict their total and priced risk: Effects of bank technical sophistication and learning over time. *Review of Accounting Studies* 9 (2–3): 265–294.
- Liu, C., and S. Ryan. 2006. Income smoothing over the business cycle: Changes in banks' coordinated management of provisions for loan losses and loan charge-offs from the pre-1990 bust to the 1990s boom. *The Accounting Review* 81 (2): 421–441.
- Mayers, D., and C. W. Smith. 1982. On the corporate demand for insurance. *The Journal of Business* 55 (2): 281–296.
- McAnally, M. L. 1996. Banks, risk, and FAS 105 disclosures. *Journal of Accounting, Auditing & Finance* 11 (3): 453–490.
- Mian, S. 1996. Evidence on corporate hedging policy. *Journal of Financial and Quantitative Analysis* 31 (3): 419–439.
- Montesi, C., and T. Lucas. 1996. Toward improved accounting for derivatives and hedging activities. *The CPA Journal* 66 (10): 20–26.
- Morgan, D. P., and K. J. Stiroh. 2001. Market discipline of banks: The asset test. *Journal of Financial Services Research* 20 (2–3): 195–208.
- Nance, D. R., C. W. Smith, and C. W. Smithson. 1993. On the determinants of corporate hedging. *The Journal of Finance* 48 (1): 267–284.
- Nelson, K. 1996. Fair value accounting for commercial banks: An empirical analysis of SFAS No. 107. *The Accounting Review* 71 (2): 161–183.
- Nelson, C. R., and A. F. Siegel. 1987. Parsimonious modeling of yield curves. *The Journal of Business* 60 (4): 473–489.
- Office of the Comptroller of Currency. 2005. *Interagency Advisory on Accounting and Reporting for Commitments to Originate and Sell Mortgage Loans*. OCC 2005-18. Washington, D.C.: Government Printing Office.
- Puwalski, A. C. 2003. *Derivatives Risk in Commercial Banking*. Washington, D.C.: Financial Analysis Section of the Division of Insurance and Research, Federal Deposit Insurance Company.

- Rane, D. 1992. A hedge or not a hedge? EITF 91-4 tried to answer the question. *Journal of Corporate Accounting & Finance* 3 (3): 279-284.
- Rogers, D. A. 2002. Does executive portfolio structure affect risk management? CEO risk-taking incentives and corporate derivatives usage. *Journal of Banking & Finance* 26 (2-3): 271-295.
- Ryan, S. G. 2007. *Financial Instruments and Institutions: Accounting and Disclosure Rules*. New York, NY: John Wiley & Sons, Inc.
- Schrand, C. M. 1997. The association between stock-price interest rate sensitivity and disclosures about derivative instruments. *The Accounting Review* 72 (1): 87-109.
- , and H. Unal. 1998. Hedging and coordinated risk management: Evidence from thrift conversions. *The Journal of Finance* 53 (3): 979-1013.
- Sengupta, P. 1998. Corporate disclosure quality and the cost of debt. *The Accounting Review* 73 (4): 459-474.
- Shackelford, D., and T. Shevlin. 2001. Empirical tax research in accounting. *Journal of Accounting and Economics* 31 (1-3): 321-387.
- Shyu, Y., and A. K. Reichert. 2002. The determinants of derivative use by U.S. and foreign banks. *Research in Finance* 19 (1): 143-172.
- Sinkey, J. F., and D. A. Carter. 2000. Evidence on the financial characteristics of banks that do and do not use derivatives. *The Quarterly Review of Economics and Finance* 40 (4): 431-449.
- Smith, C., and R. Stulz. 1985. The determinants of firms' hedging policies. *Journal of Financial and Quantitative Analysis* 20 (4): 391-405.
- Stulz, R. M. 1984. Optimal hedging policies. *Journal of Financial and Quantitative Analysis* 19 (2): 127-140.
- Swad, S. M. 1995. Accounting and disclosures for derivatives. Presentation at the 1995 Twenty-Second Annual National Conference on Current SEC Developments, Office of the Chief Accountant, U.S. Securities and Exchange Commission, Washington, D.C.
- Tufano, P. 1996. Who manages risk? An empirical examination of risk management practices in the gold mining industry. *The Journal of Finance* 51 (4): 1097-1137.
- Venkatachalam, M. 1996. Value relevance of banks' derivative disclosures. *Journal of Accounting and Economics* 22 (1-3): 327-355.
- Wahlen, J. 1994. The nature of information in commercial bank loan loss disclosures. *The Accounting Review* 69 (3): 455-478.
- Wong, M. H. F. 2000. The association between SFAS No. 119 derivatives disclosures and the foreign exchange risk exposure of manufacturing firms. *Journal of Accounting Research* 38 (2): 387-417.

Supply Chains and Segment Profitability: How Input Pricing Creates a Latent Cross-Segment Subsidy

Anil Arya
Brian Mittendorf
The Ohio State University

ABSTRACT: Recent years have seen an increased emphasis on developing more precise accounting measures of market- and customer-level profitability. By their very nature, such segment profitability measures may unwittingly neglect complementarities. Such complementarities have been widely recognized on the demand side due to brand recognition, predatory pricing, or interdependent products. We develop a model showing that important supply-side complementarities can also be prevalent. In particular, when a firm relies on a self-interested supplier for inputs used across multiple segments, the wholesale price it pays depends on the average profitability of its segments. Taking such interaction between upstream pricing and the firm's downstream reach into account, the model shows that: (1) segment profit calculations can understate or overstate the value added by the segment depending on the segment's relative contribution margin, and (2) the firm sometimes benefits from devoting resources to less profitable segments and perhaps even from serving seemingly unprofitable markets and/or customers.

Keywords: *input pricing; performance measurement; segment profit; supply chains.*

I. INTRODUCTION

Accountants have been persistent in attempts to generate disaggregate data that provides meaningful measures of segment profitability at both the market and customer levels. While joint costs and resource utilization have long been sticking points in such endeavors, recent developments stemming from activity-based costing (ABC) and customer lifetime value (CLV) analysis have been highly touted as means of better determining which markets and customers to service and the level of resources to devote to them. As a result, reliance on such measures has grown substantially in recent years.

We thank David Aboody, Mark Bagnoli, Anne Beyer, John Christensen, Peter Christensen, Sunil Dutta, Harry Evans (editor), John Fellingham, Hans Frimor, John Hughes, Steve Kachelmeier (senior editor), James Peck, Thomas Pfeiffer, Suresh Radhakrishnan, Madhav Rajan, Florin Sabac, David Sappington, Susan Watts, Lixin Ye, seminar participants at the University of Alberta, University of California, Los Angeles, The Ohio State University, Purdue University, The University of Texas at Dallas, University of Toronto, University of Vienna, and two anonymous referees for helpful comments. Professor Arya gratefully acknowledges support from the John J. Gerlach Chair.

Editor's note: Accepted by John H. (Harry) Evans III.

Submitted: December 2009
Accepted: October 2010
Published Online: April 2011

While the desire to finely parse performance is understandable, history is rife with examples in which refined performance measures failed to fully reflect opportunity costs of decisions.¹ In the context of market- and customer-level profitability, such opportunity costs have been recognized in terms of demand-side interdependencies. As examples, the benefits of having a loss-leader product for complementary demand-side effects on other products (e.g., “freebie marketing”) or incurring sustained losses for competitive posturing (e.g., dumping or predatory pricing) have been well recognized. As such, practitioners and academics alike are acutely aware of demand-side considerations when making resource allocation choices.

In this study, we demonstrate that, in addition to the well-documented demand-side spillovers, supply-side spillovers can also be present, necessitating adjustments to segment profits in order to evaluate the value added by a segment. In particular, even if market and customer profitability measures closely track individual segment profitability, they do not fully reflect the cross-subsidization that can arise when a multi-segment firm relies on a supplier for key inputs. The cross-subsidization stems from the fact that supplier pricing is influenced by the extent and profitability of the buying firm’s overall downstream reach.

To elaborate, the input price secured by a conglomerate depends on the conditions faced by the portfolio of markets in which it operates. As a result, altering the portfolio of markets that a firm serves alters the supply environment as well. A large retail conglomerate that has expanded beyond the “low-hanging fruit” of its prime markets to markets with more modest profitability can secure lower input prices. Conventional wisdom suggests such lower prices arise due to the buyer power that greater size affords. The results here suggest that there is more to buyer power than mere size; when a firm procures inputs for multiple uses, input prices are affected by average market conditions, and thus expansion to less profitable markets has the concomitant silver lining of lowering the prevailing input price.²

The basic model is of a firm that relies on a supplier for an input of material or labor that is used across multiple markets reflecting different segments or customers. The markets can differ in both the extent of retail demand and the degree of competition. With a presence in multiple markets, the seller is charged a wholesale price for the inputs it procures. Because the price does not dictate how or where the item has to be used, the firm is able to effectively prevent the input supplier from price discriminating based on usage. As such, when evaluating a segment’s worth to the firm, the critical issue is not only an individual market’s profitability, but also that market’s effect on the firm’s purchase sensitivity to input pricing.

Having identified this supply-side consideration, we next examine the circumstances under which such forces arise. In particular, given that the segments differ in both competition and retail demand, we identify the underlying drivers of cross-segment complementarities due to their effect on input price sensitivity. Interestingly, determining a segment’s net effect on the firm’s input price, and hence the direction of its supply-side effect, amounts to a concise comparison of a competition-weighted contribution margin for that segment to the competition-weighted average contribution margin of the other segments.

In this realm, practitioners routinely criticize firms for empire building and apparent corporate socialism, in that a firm that is highly successful in one arena tends to expand into other arenas

¹ Commonly discussed examples include the fact that accounting accruals occur in undiscounted terms and thus fail to reflect the cost of capital, product line costs do not reflect constrained capacity, and fixed cost allocations potentially induce a “death spiral” (see, for example, Zimmerman 2003).

² Consistent with this view, in recent times, major American auto makers have argued that their large dealer networks to rural and less-profitable metropolitan areas (as compared to the small, targeted dealership network of their foreign counterparts) make them deserving of lower prices from parts suppliers, more favorable labor contracts, and even cheaper capital (from government loans).

with less success. As an example, Toyota was widely panned for its push of the Lexus brand in Japan where the luxury vehicle market is dominated by European brands (e.g., Chozick 2007). Even if this effort does not yield the demand-side benefits of boosted sales in its home market, it can nonetheless help provide downward pressure on input prices that also affect its core business of luxury car sales in the U.S. In a similar vein, the slow start in U.S. sales of the new Acura ZDX does not bode as poorly for the firm as one may think, since the ZDX relies on as many common parts as the more widely profitable Acura MDX and Honda Pilot models and may, thus, help keep input prices low on these more profitable models.

Put differently, if the firm chooses to eliminate an underperforming segment and operate only its most successful product line, firm-wide profit would decrease by more-than-indicated segment profit figures alone due to the fact that the low-profit segment plays a large role in restricting the supplier wholesale price markup. Because such wholesale price adjustments are off-equilibrium, they are not reflected in (on-equilibrium) profit calculations and, as such, represent a latent subsidy provided by less-profitable markets to more-profitable markets.

As we demonstrate, this force means that it can even be the case that expanding into a seemingly unprofitable market may be optimal for a profit-maximizing firm. That is, when expanding (1) yields modestly negative profit in the new segment and (2) such losses are partly due to low margins of the new segment, it can be the case that expansion is worthwhile due to supply-side complementarities despite the negative segment profit that ensues. Our analysis provides precise conditions under which this phenomenon can (and cannot) occur. The theme of the analysis is further explored by examining the consequences of rivals' reliance on the same supplier as the multi-segment firm, and considering the effects of price competition in the retail realm. We demonstrate that this basic theme is robust with respect to these variations, though subtle new considerations arise.

The thrust of this study lies at the intersection of discussions about (1) measuring segment profitability and (2) coordinating self-interested supply chain partners. In the realm of profitability measurement, activity-based costing (ABC) has created a substantial improvement in the ability to track profits at the market or customer level (e.g., Cooper and Kaplan 1988). Through improved cost allocations, these processes are viewed by both academics and practitioners as critical to properly evaluating product line choices, resource allocation across segments, and keep-or-drop decisions (e.g., Goebel et al. 1998; Searcy 2004; Sopariwala 2005). At the customer level, ABC and customer lifetime value (CLV) calculations are used jointly to evaluate which customers to serve and target (e.g., Hogan et al. 2002; Searcy 2004; Venkatesan and Kumar 2004; Kuchta and Troska 2007). While significant efforts have been made to incorporate effects of downstream competition and demand-side complementarities in such calculations, our results point to strategic upstream supplier pricing as another issue worthy of consideration.

Supply chain frictions have been scrutinized by researchers in many areas of management. A fundamental concern at the forefront of such studies is the inherent conflict of interest in pricing. While supply chain partners have many common goals, individual interests can lead to inefficiencies that harm all. Perhaps most commonly analyzed among supply chain inefficiencies is that arising from double-marginalization. The issue of double-marginalization was first analyzed by Spengler (1950), following which several studies have considered the extent of pricing-induced distortions in supply chains and the myriad of means to alleviate such distortions (for summaries, see Katz (1989) and Lariviere [2008]).³ We examine whether and how supply chain coordination concerns due to double marginalization interact with decisions a seller makes with regard to the

³ Research in this area has also examined circumstances where prices above cost (and double-marginalization) prove helpful. Explanations include strategic customer behavior (Su and Zhang 2008), the time-inconsistency problem and

markets it serves. In doing so, we demonstrate a supply-chain-induced interdependency among product lines and customers in a conglomerate that is not reflected in standard segment profit calculations.

Extant research has also stressed other forms of supply chain inefficiency, manifest in terms of moral hazard among supply chain participants (e.g., Baiman et al. 2000; Balachandran and Radhakrishnan 2005; Hwang et al. 2006; Saouma 2008), adverse selection due to private information of supply chain members (e.g., Baiman and Rajan 2002), as well as a combination of these frictions (e.g., Baiman et al. 2001). In this vein, we stress that while our analysis examines a key interdependency among segments due to supply chain frictions in the context of double-marginalization, the key insight is not from double-marginalization *per se* but instead in showing that nontrivial supply chain frictions can induce supply-side spillovers that are not fully appreciated in segment profit calculations. Consistent with this point, we note that similar results can be obtained if the source of supply chain friction is due to adverse selection instead of double-marginalization.⁴

II. MODEL

A downstream (retail) firm relies on an upstream (wholesale) supplier for a key input. In converting the input and providing final goods to consumers, the downstream provider can reach multiple markets, some of which are more profitable than others. In particular, the retailer serves n segments, $n \geq 2$, and requires one unit of input per unit of output in each segment. In segment i , $i = 1, \dots, n$, the retailer faces competition from a (Cournot) competitor, denoted rival i . For now, presume rivals do not rely on the inputs provided by the supplier. Retail demand in segment i for the firm (rival) is reflected by the inverse demand curve $p_i = a_i - q_i - \gamma_i \tilde{q}_i$ ($\tilde{p}_i = a_i - \tilde{q}_i - \gamma_i q_i$), where p_i ($\tilde{p}_i$) is the firm's (rival's) retail price for each unit, and q_i ($\tilde{q}_i$) is the firm's (rival's) number of units sold. In the demand curve, $a_i, a_i > 0$, captures the degree of retail demand, and γ_i , $0 \leq \gamma_i \leq 1$ captures the degree of competitive intensity.⁵

Given these retail market conditions, the supplier determines the wholesale price for its inputs, w . For simplicity, we normalize the supplier's production cost and the firm's and rivals' other costs to zero. Upon observing the prevailing wholesale price, the firm and its rivals (concurrently) choose the quantity of output to provide to their respective retail markets. Denote the supplier's chosen wholesale price by w^n and the competitive retail outcome by $q^n = \{q_i^n\}_{i=1}^n$ and $\tilde{q}^n = \{\tilde{q}_i^n\}_{i=1}^n$ (in each case, the superscript n denotes the equilibrium when the firm participates in n markets). Given this outcome, segment i profit, $i = 1, \dots, n$, is $\pi_i(q_i^n, \tilde{q}_i^n, w^n) = [a_i - q_i^n - \gamma_i \tilde{q}_i^n]q_i^n - w^n q_i^n$, and firm profit is $\Pi^n(q^n, \tilde{q}^n, w^n) = \sum_{i=1}^n \pi_i(q_i^n, \tilde{q}_i^n, w^n)$.

Given this setup, our interest is in determining the value added by a segment. Without loss of generality, consider the value of segment n . Using the above notation, in equilibrium, segment n 's realized accounting profit is $\pi_n(q_n^n, \tilde{q}_n^n, w^n) \equiv \pi_n^*$. The value added by segment n is firm profit with segment n less firm profit that would be realized in the absence of segment n . That is, the value added by segment n , denoted V_n , is $V_n = \Pi^n(q^n, \tilde{q}^n, w^n) - \Pi^{n-1}(q^{n-1}, \tilde{q}^{n-1}, w^{n-1})$. The question we address in subsequent analyses is if and how accounting segment profit, π_n^* , differs from segment value added, V_n .

III. RESULTS

Equilibrium Supply Contracts

Before comparing the accounting measure of segment value with the underlying value added by a segment, we first derive the equilibrium outcome for operations in a given set of segments.

4 durable goods (Arya and Mittendorf 2006), and the desire to undercut *de facto* alliances between a supplier and a rival (Arya et al. 2008).

5 A previous version of this study provided details of such an analysis.

6 Throughout the analysis, we presume the parameters are such that there is nontrivial participation in each segment in equilibrium.

Then, comparing outcomes under n versus $n-1$ segments can provide a full characterization of the effects of segment n on firm profit. Working backward in the game, given the quoted wholesale price, w , and each rival's presumed retail quantity, $\tilde{q}_i$, the firm chooses retail quantities to maximize its total profit as in (1):

$$\text{Max}_{q_i, i=1, \dots, n} \sum_{i=1}^n [a_i - q_i - \gamma_i \tilde{q}_i - w] q_i. \quad (1)$$

Similarly, each rival chooses $\tilde{q}_i$ to maximize its retail profit given q_i :

$$\text{Max}_{\tilde{q}_i} [a_i - \tilde{q}_i - \gamma_i q_i] \tilde{q}_i. \quad (2)$$

Jointly solving the first-order conditions of (1) and (2) yields:

$$q_i(w) = \frac{a_i[2 - \gamma_i] - 2w}{4 - \gamma_i^2} \quad \text{and} \quad \tilde{q}_i(w) = \frac{a_i[2 - \gamma_i] + \gamma_i w}{4 - \gamma_i^2}.$$

Intuitively, the higher the firm's marginal cost (w), the lower is its quantity and the greater is its rival's quantity. This feature gives rise to the supplier's induced demand, which governs the supplier's problem. In particular, the supplier chooses a wholesale price that maximizes its profit realized from the firm; this profit is derived from all n retail segments of the firm, as follows:

$$\text{Max}_w \sum_{i=1}^n w q_i(w) \Leftrightarrow \text{Max}_w \sum_{i=1}^n w \left[\frac{a_i[2 - \gamma_i] - 2w}{4 - \gamma_i^2} \right]. \quad (3)$$

The first-order condition of (3) is $\sum_{i=1}^n \frac{a_i[2 - \gamma_i] - 4w}{4 - \gamma_i^2} = 0$. Denoting $k_i = \frac{1}{4 - \gamma_i^2}$ and $\alpha_i = \frac{a_i}{2 + \gamma_i}$ and the corresponding average values by $\bar{k}^n$ and $\bar{\alpha}^n$ (i.e., $\bar{k}^n = [1/n] \sum_{i=1}^n k_i$ and $\bar{\alpha}^n = [1/n] \sum_{i=1}^n \alpha_i$), the first-order condition yields $w^n = \bar{\alpha}^n / [4\bar{k}^n]$. Intuitively, the supplier's chosen wholesale price is affected by two key factors: the level of retail demand and the degree of competition. The greater the retail demand (the greater α_i), the greater is the firm's willingness to pay and, thus, the greater is the wholesale price. The greater the competitive intensity (the greater k_i), the greater is the threat that the rival will poach retail territory (and, hence, the supplier's wholesale territory) in response to a higher wholesale price. As a result, more competitive intensity translates into a lower chosen wholesale price. When this thinking is expanded to a firm serving n distinct retail markets, the supplier's price is similarly affected, but now the wholesale price reflects the cross-segment averages of each key factor ($\bar{\alpha}^n$ and $\bar{k}^n$).

Using this equilibrium wholesale price in the firm's and its rivals' quantity choices reveals equilibrium retail quantities, $q_i(w^n)$ and $\tilde{q}_i(w^n)$. Lemma 1 summarizes the outcome of this exercise. (All proofs are in the Appendix.)

Lemma 1: When the firm operates in n segments, the equilibrium outcome entails:

- (i) $q_i^n = \frac{1}{2} \left[2\alpha_i - \frac{k_i}{\bar{k}^n} \bar{\alpha}^n \right]$, $\tilde{q}_i^n = \alpha_i + \frac{\gamma_i k_i}{4} \bar{\alpha}^n$, and
- (ii) $w^n = \frac{\bar{\alpha}^n}{4\bar{k}^n}$.

The lemma provides a crisp presentation of the above intuition. In Lemma 1(ii), as emphasized above, the supplier's wholesale price reflects a weighing of the retail demand and retail competition effects on wholesale demand. In each case, the relevant parameter reflects an average of those parameters that underlie each segment. The resulting procurement quantities reflect a

similar averaging. The quantity chosen by the firm in a given segment is increasing in that segment's demand (α_i) relative to the average demand of its segments ($\bar{\alpha}^n$). This latter effect means that the average demand of all segments, by boosting the wholesale price, serves to moderate the quantity choice of the firm in any given segment.

Similarly, the greater the competitive intensity a firm faces in any given segment (k_i) relative to its average intensity ($\bar{k}^n$), the lower is its quantity due to its rival grabbing market share. This consideration also has reverberations to the rivals' choices in that the firm's average market characteristics, by affecting wholesale prices and, in turn, the firm's retail choices, alter rival quantities. Thus, in terms of both a firm's competitive choices and its rivals' choices, a cross-segment interdependency arises solely due to upstream pricing considerations. In the case of the rival, this interdependency arises despite the fact that the rival does not rely on the supplier and participates in a single market. Given the equilibrium outcomes, we next consider if and how segment profit differs from segment value added.

Segment Value Added versus Segment Profit

Segment profit is the primary means through which a firm and outsiders view the contribution each segment makes to the firm's overall profit. When evaluating a segment's worth, it is often discussed that accounting profit measures can leave out positive or negative demand-side spillovers across segments. In particular, one product line may be a "loss leader" in that its segment profit is negative, yet it is worth maintaining due to its positive effect on the demand for other, more profitable, product lines. Inkjet printers represent a prominent example—printers are often sold near or below cost, but create a captive customer for cartridges that are sold at a substantial profit. In that case, looking at poor margins on printers reflected in segment profit clearly does not represent the whole story.

In a similar manner to such demand-side effects, this study seeks to identify the role of supply-side effects on segment value added. In both cases, accounting segment profit figures fairly reflect "what is," but may fail to answer "what if." Our measure of segment value added, on the other hand, is a measure of the incremental value of a segment, in that it compares current firm value with the value that would have been attained in the absence of that segment. As the following lemma demonstrates, the difference between segment profit and segment value added can be captured by supply-side and demand-side spillovers.

Lemma 2: Segment n 's value added, V_n , can be expressed as:

$$V_n = \pi_n^* + S_n + D_n, \text{ where}$$

$$S_n = \Pi^{n-1}(q^n, \tilde{q}^n, w^n) - \Pi^{n-1}(q^n, \tilde{q}^n, w^{n-1}), \text{ and}$$

$$D_n = \Pi^{n-1}(q^n, \tilde{q}^n, w^{n-1}) - \Pi^{n-1}(q^{n-1}, \tilde{q}^{n-1}, w^{n-1}).$$

As Lemma 2 demonstrates, segment value added differs from segment profit due to supply-side spillovers (S_n) and demand-side spillovers (D_n). In particular, S_n reflects the change in profit of the $n-1$ remaining segments that would be induced by the concomitant change in supplier pricing, holding constant quantities demanded. Similarly, D_n reflects the change in profit of the $n-1$ remaining segments that would be induced by the concomitant change in retail quantities (demand), holding constant the wholesale price. While demand-side spillovers play a prominent role in the lexicon of business strategy as evidenced by the prominence in textbooks, academic research, and practitioner literature, the role of supply-side spillovers has not been demonstrated. The present analysis addresses this gap.

With this in mind, the most pertinent expression for the present analysis is S_n , which can be written more simply as $S_n = [w^{n-1} - w^n][q_1^n + \dots + q_{n-1}^n]$. Thus, the sign of S_n amounts to determining whether segment n 's presence serves to increase or decrease the prevailing wholesale price. If $S_n > 0$, then the wholesale price is reduced by operation in segment n ; in such a case, segment n provides a latent supply-side subsidy to the other segments.

Determination of the extent of both the supply-side and demand-side effects requires substituting the equilibrium values q^n , $\bar{q}^n$, q^{n-1} , $\bar{q}^{n-1}$, w^n , and w^{n-1} from Lemma 1 into $\Pi^{n-1}(q, \bar{q}, w)$. This exercise reveals:

$$S_n = \left[\frac{k_n}{4(n-1)\bar{k}^{n-1}} \right] \left[\frac{\bar{\alpha}^n}{\bar{k}^n} - \frac{\alpha_n}{k_n} \right] \left[\sum_{i=1}^{n-1} q_i^n \right], \text{ and}$$

$$D_n = - \left[\frac{k_n^2}{4(n-1)\bar{k}^{n-1}} \right] \left[\frac{\bar{\alpha}^n}{\bar{k}^n} - \frac{\alpha_n}{k_n} \right] \left[\left(\frac{\bar{\alpha}^n}{\bar{k}^n} - \frac{\alpha_n}{k_n} \right) \left(\frac{\sum_{i=1}^{n-1} k_i^2}{(n-1)\bar{k}^{n-1}} \right) - \frac{\sum_{i=1}^{n-1} k_i \gamma_i^2 q_i^n}{k_n} \right].$$

Summing S_n and D_n yields the following proposition:

Proposition 1: Segment n 's profit understates the value it provides if it is weaker than average (i.e., $\frac{\alpha_n}{k_n} < \frac{\bar{\alpha}^n}{\bar{k}^n}$) and overstates value if it is stronger than average (i.e., $\frac{\alpha_n}{k_n} > \frac{\bar{\alpha}^n}{\bar{k}^n}$).

Two implications follow from the S_n and D_n expressions that underlie the proposition. First, the supply-side effect can either be positive or negative, depending on how the characteristics of the segment in question differ from that of other segments. Second, the demand-side effect is unequivocally negative. That is, this setting is one where the typical demand-side complementarity is not at play. Instead, if a complementarity is to arise, it is on the supply side. Importantly, this supply-side effect is of first-order importance, as evidenced by the fact that the condition that determines its sign is also the condition that determines the sign of the net effect, as confirmed in Proposition 1.

Proposition 1 completes the comparison of value added and segment profit. To be more precise, it confirms that $V_n - \pi_n^* = S_n + D_n = \lambda \left[\frac{\bar{\alpha}^n}{\bar{k}^n} - \frac{\alpha_n}{k_n} \right]$, $\lambda > 0$. In other words, despite the $2n$ parameters that underlie the setup ($\alpha_1 \dots \alpha_n$ and $k_1 \dots k_n$), the nature of the latent interdependency introduced by a segment can be succinctly captured by four parameters (α_n , $\bar{\alpha}^n$, k_n , and $\bar{k}^n$). To get a better intuitive feel for the condition in the proposition, we next isolate cross-segment variability in demand and competition, each in turn.

First, say the degree of competition is the same among all retail markets—i.e., $\gamma_1 = \dots = \gamma_n$. In this case, with $k_n = \bar{k}^n$, the condition under which segment profit differs from value added amounts to a comparison of demand in that segment to the average demand among the portfolio of segments. In particular, segments with lower demand serve to indirectly subsidize those with higher demand due to their ability to moderate the prevailing wholesale price. This subsidy is captured by the segment's value added, but it is not reflected in the segment's profit number.

Corollary 1: In the $\gamma_1 = \dots = \gamma_n$ case, segments with demand that is below (above) the average demand have profits that understate (overstate) the value they provide. That is, if $\alpha_n < \bar{\alpha}^n$, then $\pi_n^* < V_n$, and if $\alpha_n > \bar{\alpha}^n$, then $\pi_n^* > V_n$.

On the other hand, if demand is the same across segments—i.e., $\alpha_1 = \dots = \alpha_n$, the only cross-market differences are reflected in varying degrees of competition. Note that the condition in the

proposition depends on α_n and $\bar{\alpha}^n$, which reflect competition-weighted measures of demand. So, despite having equal demand (intercepts), cross-market variability in the α_i terms can remain due to differences in competition. In this case, the condition in the proposition simplifies to a comparison of competition in segment n to the average competition level among the other segments. However, the average is not equally weighted; instead, the weights depend on the cross-segment variability in competition. This result is presented in Corollary 2.

Corollary 2: In the $a_1, \dots, a_n$ case, segments with competition that is above (below) a weighted-average competition level have profits that understate (overstate) the value they provide. That is, if $\gamma_n > \sum_{i=1}^n \left[\frac{k_i}{n k^n} \right] \gamma_i$, then $\pi_n^* < V_n$, and if $\gamma_n < \sum_{i=1}^n \left[\frac{k_i}{n k^n} \right] \gamma_i$, then $\pi_n^* > V_n$.

The intuition for the results in Corollary 2 can be seen by examining the case of two segments. From the corollary, in the two-segment case with $a_1 = a_2$, profit in segment 2 understates value added if and only if $\gamma_2 > \gamma_1$. In effect, with $\gamma_2 > \gamma_1$, segment 2 provides a supply-side subsidy that is not picked up by accounting profit calculations, because competitive pressures in segment 2 help drive down the wholesale price that governs activity in both segments.

A natural follow-up question to the above analysis is how one could expect to see these parameters manifest in practice. While one may get a feel for the relative competitiveness of an industry or segment with ease, the notion of a higher or lower demand intercept is perhaps less natural. However, if one instead looks at profit margins, a related picture arises. In particular, denote the (unit) contribution margin of segment i by m_i , $m_i = a_i - q_i^n - \gamma_i q_i^n - w^n$, and the n -segment average by $\bar{m}^n$, $\bar{m}^n = [1/n] \sum_{i=1}^n m_i$. Given this notation, the next proposition confirms that the succinct comparison in Proposition 1 also has an accounting analog.

Proposition 2: Segment n 's profit understates the value it provides if its profit margin is weaker than average (i.e., $\frac{m_n}{k_n} < \frac{\bar{m}^n}{k^n}$) and overstates value if its profit margin is stronger than average (i.e., $\frac{m_n}{k_n} > \frac{\bar{m}^n}{k^n}$).

In short, the proposition shows that a segment that faces greater competition and/or lower contribution margins is one that serves an important off-equilibrium role in boosting the fortunes of other segments. Besides demonstrating the underlying forces driving an inherent inter-linkage in segments when a firm relies on external input supply, the proposition may shed some light on the widely held view that in multi-segment firms, underperforming segments are propped up by overperforming segments (e.g., Shin and Stulz 1998; Rajan et al. 2000; Gertner et al. 2002). Since such findings are derived in part from accounting segment profit measures, they may not reflect the latent input price subsidy provided in the reverse direction—i.e., a subsidy to overperforming segments provided by an underperforming segment.

Even more broadly, the proposition suggests that a firm with multiple segments that utilize common inputs would exhibit seemingly excessive (constrained) investment in underperforming (overperforming) segments, despite such behavior being value-maximizing. Viewed in this light, apparent inefficiencies in resource allocation may not be evidence of value destruction from diversification but rather reflect a form of selection bias in that multi-segment firms are inherently different from their single-segment counterparts (see, e.g., Chevalier 2004). Roughly stated, contrasting a multi-segment firm with a linear disaggregation of its segments misses out on the fact that third-party relationships of presumptive, stand-alone segments can be strikingly different from those nurtured by the aggregate firm. Clearly, such shifts must be accounted for in order to calculate a meaningful segment diversification discount/premium. If this view were borne out empirically, then one would expect the most (least) profitable divisions of a multi-segment firm to

perform better (worse) than their single-segment counterparts. Further, this feature would be expected to be more pronounced the more critical input supply is to the multi-segment firm.

Extending the theme, we next address if and how the latent supply-side subsidy provided by underperforming segments can alter the decision to enter a market in the first place.

The Decision to Enter a Market

The analysis thus far has demonstrated how accounting measures of segment performance can fail to capture the benefits of poorly performing segments due to their ability to put downward pressure on supplier pricing. Among other things, this view can possibly shed new light on seemingly widespread corporate socialism. When one thinks of criticisms of corporate socialism, wherein a firm uses one very successful line of business to seemingly support otherwise unsustainable segments, the notions of empire building and excessive market entry in response to success in a core market inevitably come to mind. To this end, we next consider the implications of the model for decisions to enter ancillary markets.

Suppose the firm operates in markets $1, \dots, n-1$ and is considering entry into market n . Denoting the cost of entry into market i by F_i , the segment profit gleaned from entry into market n is simply $\pi_n^* - F_n$. Suppose further that this proposed expansion fits the bill for empire building in that the potential profit from the segment is below the average for the firm's core segments, i.e., $\frac{\alpha_n}{k_n} < \frac{\bar{\alpha}}{\bar{k}}$. In that case, from Proposition 1, entry into the new segment offers a latent benefit of favorably altering the prevailing wholesale price. In fact, this force, while latent, may form the basis for expansion into markets that yield negative profit. In other words, a firm may expand into new "loss-leader" markets, not as demand-side strategic posturing but instead as supply-side strategic posturing. The next proposition presents the conditions under which this phenomenon arises.

Proposition 3: If $\frac{\alpha_n}{k_n} < \frac{\bar{\alpha}}{\bar{k}}$, then the segment profit from entry is negative while the value added by entry is positive for all F_n in the non-empty interval $(F', F' + S_n + D_n)$, where $F' = \left[\alpha_n - \frac{k_n}{2\bar{k}} \bar{\alpha}^n \right]^2$.

In effect, the proposition notes that if a new segment is weaker than the existing portfolio of segments, then it is possible that entry therein yields negative segment profit but is nonetheless worthwhile. This occurs only for intermediate levels of entry costs. If the cost of entry is low ($F_n < F'$), then the positive margins from entry are sufficient to cover fixed costs, and segment profit in the new segment is positive. At the other extreme, for high entry costs ($F_n > F' + S_n + D_n$) the entry cost is so large that even when the supply-side complementarity is considered, entry is not worth the cost. For an intermediate interval of entry costs ($F_n \in (F', F' + S_n + D_n)$), the segment profit is negative but the supply-side subsidy makes entry worthwhile. Consistent with this, the larger the supply-side effect (S_n), the larger the interval.

Proposition 3 is couched in terms of implications of strategic input price adjustments on firms' entry decisions. Of course, in an analogous fashion, such concerns also impact a firm's decision to exit a market, where F_i represents the ongoing (fixed) cost of operating a segment. The example in Table 1 makes the point that the failure to recognize that the wholesale price is variable (not fixed) with respect to the decision to keep or drop a segment can lead to a death spiral. In a sense, this is the "reverse" of the familiar cost-allocation death spiral in which allocation can create an impression that a cost is variable while it really is fixed (and sunk). Broadly stated, both the supply-chain death spiral and the cost-allocation death spiral stress the need to be cognizant of the fixed versus variable cost dichotomy.

In Table 1, the first scenario characterizes the equilibrium outcomes when the firm operates in all three segments. In this case, the firm earns positive profit, but does so in the presence of a loss

TABLE 1
A Supply Chain Death Spiral

Example Parameters ($\gamma_i = 0$)

	<u>Segment 1</u>	<u>Segment 2</u>	<u>Segment 3</u>
Demand parameter (a_i)	40	30	20
Fixed cost (F_i)	110	45	10

Firm Operates in Segments 1, 2, and 3: ($w = 15$)

	<u>Segment 1</u>	<u>Segment 2</u>	<u>Segment 3</u>	<u>Total</u>
Output level (q_i)	12.50	7.50	2.50	22.50
Profits (π_i)	46.25	11.25	-3.75	53.75

Firm Operates in Segments 1 and 2: ($w = 17.5$)

	<u>Segment 1</u>	<u>Segment 2</u>	<u>Total</u>
Output level (q_i)	11.25	6.25	17.50
Profits (π_i)	16.5625	-5.9375	10.625

Scenario 3: Firm Operates in Segment 1: ($w = 20$)

	<u>Segment 1</u>	<u>Total</u>
Output level (q_i)	10	10
Profits (π_i)	-10	-10

in segment 3. The next scenario demonstrates the outcome of following a policy of exiting the unprofitable segment. In this case, the removal of segment 3 results in a lowering of firm-wide profit (from 53.75 to 10.625), in which case segment 2 arises as unprofitable. Eliminating this segment, in turn, leads to the final scenario, wherein only one segment remains and firm profit is negative. In other words, supplier pricing effects can introduce a death spiral of sorts, despite the fact that the fixed costs of operating in each segment are presumed to disappear along with the segment. Here, the death spiral arises because elimination of unprofitable segments serves to embolden the supplier in its pricing, thereby putting downward pressure on the margins of the more profitable segments. In effect, the firm's decision of eliminating a segment is not made in isolation in that the supplier, too, reacts strategically in response to such a course.

Rivals' Reliance on the Supplier

To this point, we have presumed a setting wherein a firm relies on the supplier for inputs but its rivals are self-sustaining. Suppose instead that the rivals, too, rely on the supplier for inputs. By providing a perspective on alternative supply-market conditions and the consequences for supply-side externalities from multi-market participation, this setting can serve as a robustness check on our underlying premise. In this case, the supplier sets $n+1$ wholesale prices, one for each firm (denoted w for the multi-segment firm and w_i for rival i). As before, the firm and its rivals concurrently choose retail quantities subsequent to observation of the prevailing wholesale prices.

Lemma 3 presents the equilibrium outcome under the setting with both the firm and its rivals reliant on the supplier, where the notation $[']$ distinguishes this equilibrium from that of Lemma 1.

Lemma 3: If the firm and the rivals each rely on the supplier, then the equilibrium outcome entails:

$$(i) \quad q_i'^n = \frac{1}{2} \left[\frac{a_i}{2+\gamma_i} - \frac{\bar{a}^n - a_i}{2} \right], \quad \tilde{q}_i'^n = \frac{a_i}{2[2+\gamma_i]}, \text{ and}$$

$$(ii) \quad w'^n = \frac{\bar{a}^n}{2}, \quad \tilde{w}_i'^n = \frac{1}{2} \left[a_i + \frac{\gamma_i(\bar{a}^n - a_i)}{2} \right].$$

Comparing Lemmas 1 and 3, a key difference is apparent—whereas both the level of retail demand and the degree of retail competition played critical roles in the multi-segment firm's wholesale price in the former case, the wholesale price is independent of the degree of competition in the latter. That is, when the supplier provides inputs both to the firm and its rival, it does not tailor the firm's input price to the level of competition they face. To explain, when the supplier provides inputs only to the firm, it is that firm's *de facto* partner, and, as such, is eager to cut wholesale price to give the firm an edge in competition, i.e., as γ_i increases, w^n decreases. In contrast, when the rival is also the supplier's customer, the supplier has no incentive to favor one customer over another. The end result is that w'^n is independent of γ_i .

The desire to have a level playing field among competing firms extends to the wholesale price set for rival i , in that the rival's price reflects the average retail market conditions of the firm, the extent of which is dictated by the level of competition, γ_i . The effect of cross-segment averages extends to the rivals despite the fact that each participates in only one market. This feature arises so that if the firm benefits from an underperforming segment by getting price cuts, some of those cuts are carried over to the rival so as to maintain competition and thus boost input purchases in that market.

Given the induced wholesale prices and resulting competitive outcomes in the case of rival reliance on the supplier, the supply-side effect of participation in segment n is:

$$S'_n = \left[\frac{1}{2(n-1)} \right] [\bar{a}^n - a_n] \left[\sum_{i=1}^{n-1} q_i'^n \right],$$

whereas the demand-side effect is:

$$D'_n = - \left[\frac{1}{16(n-1)} \right] [\bar{a}^n - a_n]^2.$$

Given these expressions, the next proposition demonstrates the net effect of the firm's multi-market participation.

Proposition 4: With rival reliance on the supplier, segment n 's profit understates the value it provides if its demand is weaker than average (i.e., $a_n < \bar{a}^n$) and overstates value if its demand is stronger than average (i.e., $a_n > \bar{a}^n$).

Comparing the expressions that underlie Proposition 4 and Proposition 1, it is readily seen that the rivals' reliance on the supplier simplifies the expression for both the demand- and supply-side effects. In particular, no longer is the degree of competition in any of the markets critical, but instead the distinguishing factor is the comparison of retail demand in one segment to the average retail demand levels. This factor carries forward to the determination of whether profit over- or understates segment value added. In particular, the difference between segment value added (V'_n) and segment profit ($\pi_n'^*$) can be expressed as $V'_n - \pi_n'^* = S'_n + D'_n = \lambda' [\bar{a}^n - a_n]$, $\lambda' > 0$. This simplified expression means that a concise comparison of the demand level in segment n to the equally weighted average of that in other segments determines the sign of the difference between value added and profit.

Thus, besides demonstrating that the underlying premise herein is robust to who relies on the supplier for inputs, the setting also provides some guidance in terms of how the forces should arise in practice. That is, while retail demand in a segment relative to average demand is consistently a key determinant of any latent supply-side subsidies across segments, the degree of retail competition is crucially tied to the extent that the retail parties and the wholesale supplier form an alliance.

While the analysis so far has considered the cases wherein either the rivals are not reliant on a supplier (as in the base model) or reliant on a common supplier (as in this section), there are also supply arrangements in which each retail party relies on its own dedicated supplier. The following example contrasts the common supplier and the dedicated supplier settings, and reinforces the notion that supply-side effects can manifest themselves in a myriad of input procurement relationships.

Suppose there are two segments with $a_1 = a_2$. Moreover, for simplicity, say $\gamma_1 = 0$ and $\gamma_2 > 0$, i.e., the firm enjoys a monopoly in the first segment, while it faces a rival, rival 2, in the second segment. If the firm and the rival rely on a common supplier, then, from Lemma 3, the wholesale price is $a/2$ for each buyer. This is the same price that the supplier would have charged the firm had it operated only in segment 1. Thus, with a common supplier and $a_1 = a_2$, there are no supply-side effects and segment profit equals segment value added.

However, suppose the firm and the rival rely on separate (dedicated) suppliers. In this case, the wholesale prices for the firm and the rival are $\frac{a[2-\gamma_2][3\gamma_2+8]}{32-5\gamma_2^2}$ and $\frac{a[2-\gamma_2][16+4\gamma_2-\gamma_2^2]}{2[32-5\gamma_2^2]}$, respectively. Notice, the firm's input price is less than $a/2$, the price the firm would have been charged by its dedicated supplier had it operated only in segment 1. Thus, with dedicated suppliers, there is again a supply-side subsidy. Also, note that the wholesale prices for both parties are decreasing in γ_2 . In effect, as segment 2 becomes more competitive, the suppliers respond by lowering the input charge to their respective clients. In fact, this reduced input-pricing effect can be so pronounced that both the firm and the rival may welcome the increased competition. In the example, the firm's profits are increasing in γ_2 for $\gamma_2 > 0.62$; the same is true for rival 2 for $\gamma_2 > 0.68$. The role of multi-segment operations is critical in deriving this result. After all, if the firm and the rival each operated only in segment 2, then each faces a wholesale price of $\frac{a[2-\gamma_2]}{4-\gamma_2}$. This price is again decreasing in γ_2 but each party would be made worse off by added competition. In contrast, the multi-segment firm gains because the increased competition in segment 2 lowers the input price that applies to purchases it makes in both segments, including the segment where it enjoys a monopoly and, thus, high profit margins. And, relatedly, when competition is sufficiently intense, rival 2 gains from the firm's ceding of market share in segment 2 in order to focus on its more profitable segment 1. Not surprisingly, then, similar results on retail parties benefiting from increased competition also apply when the multi-segment firm is the only party that relies on the supplier (as in the base model).

Price Competition

Thus far, we have considered the supply-side ramifications of a firm's participation in multiple output (retail) markets in a world of Cournot (quantity) competition. As the nature of competition has often been shown to lead to reversals of results (see, e.g., Darrough 1993; Dye 2001; Verrecchia 2001), we next consider the extent to which the key implications herein are sensitive to such variation by revisiting the analysis under Bertrand (price) competition. To reflect that the competitive outcome now arises under price competition, we denote all equilibrium outcomes in this case with a $[\hat{\cdot}]$. Following this notation, then, let $\hat{k}_i = \frac{2-\gamma_i}{[1-\gamma_i][4-\gamma_i]}$ and $\hat{\alpha}_i = \frac{a_i}{2+\gamma_i(1-\gamma_i)}$, and denote

the corresponding average values by $\bar{k}^n$ and $\bar{\alpha}^n$, respectively. Relegating the details to the Appendix, Lemma 4 presents the equilibrium outcomes in the case of price competition.

Lemma 4: Under retail price competition, the equilibrium outcome entails:

- (i) $\hat{q}_i^n = \frac{1}{2} \left[2\hat{\alpha}_i - \frac{\hat{k}_i}{\bar{k}^n} \bar{\alpha}^n \right]$, $\hat{q}_i^n = \hat{\alpha}_i + \frac{\gamma_i}{2(2-\gamma_i)} \frac{\hat{k}_i}{\bar{k}^n} \bar{\alpha}^n$, and
- (ii) $\hat{w}^n = \bar{\alpha}^n / [2\bar{k}^n]$.

Using these outcomes, the price-competition analogs to the supply-side and demand-side effects from the main model are:

$$\hat{S}_n = \left[\frac{\hat{k}_n}{2(n-1)\bar{k}^{n-1}} \right] \left[\frac{\bar{\alpha}^n}{\bar{k}^n} - \frac{\hat{\alpha}_n}{\hat{k}_n} \right] \left[\sum_{i=1}^{n-1} \hat{q}_i^n \right],$$

and

$$\hat{D}_n = - \left[\frac{\hat{k}_n^2}{2(n-1)\bar{k}^{n-1}} \right] \left[\frac{\bar{\alpha}^n}{\bar{k}^n} - \frac{\hat{\alpha}_n}{\hat{k}_n} \right] \left[\left(\frac{\bar{\alpha}^n}{\bar{k}^n} - \frac{\hat{\alpha}_n}{\hat{k}_n} \right) \left(\frac{\sum_{i=1}^{n-1} (1-\gamma_i^2)\hat{k}_i^2}{2(n-1)\bar{k}^{n-1}} \right) + \sum_{i=1}^{n-1} \frac{\gamma_i^2 \hat{q}_i^n}{(4-\gamma_i^2)\hat{k}_n} \right],$$

respectively. Given these expressions, the analog to Proposition 1 is as follows:

Proposition 5: Under price competition, segment n 's profit understates the value it provides if $\frac{\hat{\alpha}_n}{\hat{k}_n} < \frac{\bar{\alpha}^n}{\bar{k}^n}$ and overstates value if $\frac{\hat{\alpha}_n}{\hat{k}_n} > \frac{\bar{\alpha}^n}{\bar{k}^n}$.

In brief, comparing Propositions 1 and 5 confirms that the nature of competition (quantity versus price) does not alter the basic conclusions of the analysis. In fact, the basic form of both the supply- and demand-side effects of participation in segment n that are not accounted for in accounting segment metrics is invariant to the nature of competition. What changes instead is the precise calculation of the key parameters that capture retail demand (α_i versus $\hat{\alpha}_i$) and retail competition levels (k_i versus $\hat{k}_i$).

Taken together, the results throughout Section III suggest that (1) the importance and role of cross-segment supply-side subsidies induced by wholesale price effects is invariant to the nature of retail competition and the portfolio of supplier clients; and (2) the direction and magnitude of such supply-side effects can be critically affected by key characteristics of the retail markets (the degree of competition and extent of demand among the firm's portfolio of segments).

IV. CONCLUSION

Determining the contribution of various segments to the overall profit of a diversified entity is a delicate exercise. Refined costing measures have substantially improved this process and, hence, led to additional reliance on segment profit measures for resource allocation choices. Yet, because accounting profit metrics neglect opportunity costs, segment profit figures can deviate from the value added by a segment. The common concern in this vein is of downstream complementarities brought by loss-leader products or locations as means of furthering brand recognition and/or product loyalty. Our study examines a divergence between segment profit and segment value added that arises instead from upstream considerations.

We demonstrate that, even if product demand and long-term competitive posturing are not pressing considerations, a firm can nonetheless benefit from its weaker retail performers so as to

support more favorable supplier pricing terms. Since less profitable product lines or customer groups solidify a reluctance to accept markups in supplier pricing, a firm's active participation in less profitable markets can help support lower input prices. This effect, in turn, points to an additional consideration in keep-or-drop decisions.

While this fundamental phenomenon is demonstrated in a model wherein industrial structure is taken as given, further research could seek to endogenize the structure of industries and, in particular, the varied product offerings that a conglomerate would optimally seek in light of input-pricing considerations. Future study could also shed light on other forms of perceived "weakness" that could prove useful in supply chain relationships, such as public disclosure of proprietary information or even limited acquisition of relevant information.

APPENDIX

Proof of Lemma 1

Working backward, given wholesale price w and rival i 's output $\tilde{q}_i$, the firm chooses quantity q_i in segment i , $i = 1, \dots, n$, to maximize its total profit:

$$\text{Max}_{q_i, i=1, \dots, n} \sum_{i=1}^n [a_i - q_i - \gamma_i \tilde{q}_i - w] q_i. \quad (\text{A1})$$

Given q_i , rival i chooses $\tilde{q}_i$ to maximize its profit:

$$\text{Max}_{\tilde{q}_i} [a_i - \tilde{q}_i - \gamma_i q_i] \tilde{q}_i, \quad i = 1, \dots, n. \quad (\text{A2})$$

Jointly solving the first-order conditions of (A1) and (A2) yields:

$$q_i(w) = \frac{a_i[2 - \gamma_i] - 2w}{4 - \gamma_i^2} = \alpha_i - 2wk_i, \text{ and} \quad (\text{A3})$$

$$\tilde{q}_i(w) = \frac{a_i[2 - \gamma_i] + \gamma_i w}{4 - \gamma_i^2} = \alpha_i + w\gamma_i k_i, \quad i = 1, \dots, n.$$

The output reaction, $q_i(w)$ in (A3), serves as the induced demand function for the supplier, and it solves:

$$\text{Max}_w \sum_{i=1}^n w q_i(w) \Leftrightarrow \text{Max}_w \sum_{i=1}^n w [\alpha_i - 2wk_i]. \quad (\text{A4})$$

The first-order condition of (A4) yields w^n in part (ii). Using (A3), the outputs listed in part (i) are $q_i(w^n)$ and $\tilde{q}_i(w^n)$. This completes the proof of Lemma 1. ■

Proof of Lemma 2

The value added by segment n is V_n , $V_n = \Pi^n(q^n, \tilde{q}^n, w^n) - \Pi^{n-1}(q^{n-1}, \tilde{q}^{n-1}, w^{n-1})$. Segment n 's realized (accounting) profit is π_n^* , $\pi_n^* = \pi_n(q_n^n, \tilde{q}_n^n, w^n)$. Hence,

$$V_n - \pi_n^* = \Pi^n(q^n, \tilde{q}^n, w^n) - \Pi^{n-1}(q^{n-1}, \tilde{q}^{n-1}, w^{n-1}) - \pi_n(q_n^n, \tilde{q}_n^n, w^n). \quad (\text{A5})$$

Since $\Pi^n(q^n, \tilde{q}^n, w^n) = \Pi^{n-1}(q^n, \tilde{q}^n, w^n) + \pi_n(q_n^n, \tilde{q}_n^n, w^n)$, (A5) can be written as:

$$V_n - \pi_n^* = \Pi^{n-1}(q^n, \tilde{q}^n, w^n) - \Pi^{n-1}(q^{n-1}, \tilde{q}^{n-1}, w^{n-1}). \quad (\text{A6})$$

Adding and subtracting $\Pi^{n-1}(q^n, \tilde{q}^n, w^{n-1})$, (A6) can be written as:

$$V_n - \pi_n^* = [\Pi^{n-1}(q^n, \tilde{q}^n, w^n) - \Pi^{n-1}(q^n, \tilde{q}^n, w^{n-1})] + [\Pi^{n-1}(q^n, \tilde{q}^n, w^{n-1}) - \Pi^{n-1}(q^{n-1}, \tilde{q}^{n-1}, w^{n-1})]. \quad (A7)$$

The first term on the right-hand-side of (A7) is S_n and the second term is D_n . This completes the proof of Lemma 2. ■

Proof of Proposition 1

Because $\Pi^{n-1}(q, \tilde{q}, w) = \sum_{i=1}^{n-1} \pi_i(q_i, \tilde{q}_i, w)$ and $\pi_i(q_i, \tilde{q}_i, w) = [a_i - q_i - \gamma \tilde{q}_i - w]q_i$, the S_n and D_n expressions in Lemma 2 can be equivalently written as:

$$S_n = [w^{n-1} - w^n] \left[\sum_{i=1}^{n-1} q_i^n \right], \text{ and} \quad (A8)$$

$$D_n = \sum_{i=1}^{n-1} \{ [a_i - w^{n-1} - q_i^n - q_i^{n-1}] [q_i^n - q_i^{n-1}] - \gamma_i [\tilde{q}_i^n q_i^n - \tilde{q}_i^{n-1} q_i^{n-1}] \}.$$

From Lemma 1, $q_i^n = \frac{1}{2} [2\alpha_i - \frac{k_i}{\bar{k}^n} \bar{\alpha}^n]$, $q_i^{n-1} = \frac{1}{2} [2\alpha_i - \frac{k_i}{\bar{k}^{n-1}} \bar{\alpha}^{n-1}]$, $w^n = \frac{\bar{\alpha}^n}{4\bar{k}^n}$, $w^{n-1} = \frac{\bar{\alpha}^{n-1}}{4\bar{k}^{n-1}}$. Substituting these in (A8), and making use of the identities $\bar{\alpha}^n = \frac{n-1}{n} \bar{\alpha}^{n-1} + \frac{1}{n} \alpha_n$ and $\bar{k}^n = \frac{n-1}{n} \bar{k}^{n-1} + \frac{1}{n} k_n$, algebra yields the expressions for S_n and D_n provided in the text.

Using S_n and D_n :

$$V_n - \pi_n^* = \lambda \left[\frac{\bar{\alpha}^n}{\bar{k}^n} - \frac{\alpha_n}{k_n} \right], \text{ where} \quad (A9)$$

$$\lambda = \left[\frac{k_n}{4(n-1)\bar{k}^{n-1}} \right] \left[\sum_{i=1}^{n-1} k_i \gamma_i^2 q_i^n + \sum_{i=1}^{n-1} q_i^n - \left(\frac{k_n \bar{\alpha}^n}{\bar{k}^n} - \alpha_n \right) \left(\frac{\sum_{i=1}^{n-1} k_i^2}{(n-1)\bar{k}^{n-1}} \right) \right].$$

The sign of $V_n - \pi_n^*$ is the same as the sign of $\frac{\bar{\alpha}^n}{\bar{k}^n} - \frac{\alpha_n}{k_n}$ if $\lambda > 0$. Note that $\lambda > 0$ if and only if the term $T = \left[\sum_{i=1}^{n-1} k_i \gamma_i^2 q_i^n + \sum_{i=1}^{n-1} q_i^n - \left(\frac{k_n \bar{\alpha}^n}{\bar{k}^n} - \alpha_n \right) \left(\frac{\sum_{i=1}^{n-1} k_i^2}{(n-1)\bar{k}^{n-1}} \right) \right] > 0$. Because $q_i^n > 0$, it follows that if $\alpha_n > \frac{k_n \bar{\alpha}^n}{\bar{k}^n}$, $T > 0$. Thus, the proof is complete if it can be shown that $T > 0$ even when $\alpha_n < \frac{k_n \bar{\alpha}^n}{\bar{k}^n}$. Given $\alpha_n < \frac{k_n \bar{\alpha}^n}{\bar{k}^n}$, $\alpha_n > \frac{k_n \bar{\alpha}^n}{2\bar{k}^n}$ (from $q_n^n > 0$ in Lemma 1), $1/4 \leq k_i \leq 1/3$ (with the limiting values corresponding to $\gamma_i = 0$ and $\gamma_i = 1$, respectively), a lower bound on $\sum_{i=1}^{n-1} q_i^n$ and an upper bound on $\left(\frac{k_n \bar{\alpha}^n}{\bar{k}^n} - \alpha_n \right) \left(\frac{\sum_{i=1}^{n-1} k_i^2}{(n-1)\bar{k}^{n-1}} \right)$ is derived below:

$$\sum_{i=1}^{n-1} q_i^n = \sum_{i=1}^n q_i^n - q_n^n = \frac{n\bar{\alpha}^n}{2} - \alpha_n + \frac{k_n \bar{\alpha}^n}{2\bar{k}^n} > \frac{n\bar{\alpha}^n}{2} - \frac{k_n \bar{\alpha}^n}{\bar{k}^n} + \frac{k_n \bar{\alpha}^n}{2\bar{k}^n} = \frac{n\bar{\alpha}^n}{2} - \frac{k_n \bar{\alpha}^n}{2\bar{k}^n} > \frac{n\bar{\alpha}^n}{2} - \frac{[1/3]\bar{\alpha}^n}{2[1/4]}$$

$$= \frac{n\bar{\alpha}^n}{2} - \frac{2\bar{\alpha}^n}{3}. \quad (A10)$$

$$\left(\frac{k_n \bar{\alpha}^n}{\bar{k}^n} - \alpha_n \right) \left(\frac{\sum_{i=1}^{n-1} k_i^2}{(n-1)\bar{k}^{n-1}} \right) < \left(\frac{k_n \bar{\alpha}^n}{\bar{k}^n} - \frac{k_n \bar{\alpha}^n}{2\bar{k}^n} \right) \left(\frac{[1/9][n-1]}{[1/4][n-1]} \right) = \frac{2k_n \bar{\alpha}^n}{9\bar{k}^n} < \frac{2[1/3]\bar{\alpha}^n}{9[1/4]} \\ = \frac{8\bar{\alpha}^n}{27}. \quad (\text{A11})$$

From (A10) and (A11), and $n \geq 2$,

$$T > \sum_{i=1}^{n-1} k_i \gamma_i^2 q_i^n + \frac{n\bar{\alpha}^n}{2} - \frac{2\bar{\alpha}^n}{3} - \frac{8\bar{\alpha}^n}{27} > \sum_{i=1}^{n-1} k_i \gamma_i^2 q_i^n + \frac{\bar{\alpha}^n}{27} > 0. \quad (\text{A12})$$

Thus, $T > 0$ and, hence, $\lambda > 0$. This completes the proof of Proposition 1. ■

Proof of Corollary 1

In the $\gamma_1 = \dots = \gamma_n$ case, denote γ_i by γ . Thus, $\frac{\alpha_n}{k_n} = a_n[2 - \gamma]$ and $\frac{\bar{\alpha}^n}{\bar{k}^n} = \bar{a}_n[2 - \gamma]$. Using these in (A9), $V_n - \pi_n^* = \lambda[2 - \gamma][\bar{a}_n - a_n]$, $\lambda > 0$ and $0 \leq \gamma \leq 1$. This completes the proof. ■

Proof of Corollary 2

In the $a_1 = \dots = a_n$ case, denote a_i by a . Thus, $\frac{\alpha_n}{k_n} = a[2 - \gamma_n]$ and $\frac{\bar{\alpha}^n}{\bar{k}^n} = \frac{a}{n\bar{k}^n} \sum_{i=1}^n \frac{1}{2 + \gamma_i}$. Using these in (A9), $V_n - \pi_n^* = \lambda a \left(\gamma_n - \sum_{i=1}^n \left[\frac{k_i}{n\bar{k}^n} \right] \gamma_i \right)$, $\lambda > 0$ and $a > 0$. This completes the proof. ■

Proof of Proposition 2

When the firm operates in n -segments, its (per-unit) contribution margin in segment i is $m_i = a_i - q_i^n - \gamma_i \bar{q}_i^n - w^n$. Using outputs and wholesale price in Lemma 1, and substituting a_i by $\alpha_i[2 + \gamma_i]$ and γ_i^2 by $4 - 1/k_i$, m_i is:

$$m_i = \alpha_i - \frac{k_i}{2\bar{k}^n} \bar{\alpha}^n. \quad (\text{A13})$$

From (A13), $\bar{m}^n = \frac{1}{n} \sum_{i=1}^n m_i = \frac{\bar{\alpha}^n}{2}$. Thus,

$$\frac{\bar{m}^n}{\bar{k}^n} = \frac{\bar{\alpha}^n}{2\bar{k}^n} \quad \text{and} \quad \frac{m_n}{k_n} = \frac{\alpha_n}{k_n} - \frac{\bar{\alpha}^n}{2\bar{k}^n}. \quad (\text{A14})$$

From (A14), $\frac{\bar{m}^n}{\bar{k}^n} - \frac{m_n}{k_n} = \frac{\bar{\alpha}^n}{\bar{k}^n} - \frac{\alpha_n}{k_n}$. Thus, from (A9), $V_n - \pi_n^* = \lambda \left[\frac{\bar{m}^n}{\bar{k}^n} - \frac{m_n}{k_n} \right]$, $\lambda > 0$. This completes the proof of Proposition 2. ■

Proof of Proposition 3

Using Lemma 1 and (A13), segment n 's profit from entry is:

$$\pi_n^* - F_n = m_n q_n^n - F_n = \left[\alpha_n - \frac{k_n}{2\bar{k}^n} \bar{\alpha}^n \right]^2 - F_n = F' - F_n. \quad (\text{A15})$$

From (A15), $F_n > F'$ ensures $\pi_n^* - F_n < 0$. The value added by entry equals:

$$V_n - F_n = \pi_n^* + S_n + D_n - F_n = F' + S_n + D_n - F_n. \quad (\text{A16})$$

From (A16), the firm prefers to enter if $F_n < F' + S_n + D_n$. Thus, for $F_n \in (F', F' + S_n + D_n)$, the segment profit from entry is negative while the value added by entry is positive. Finally, from

Lemma 2 and Proposition 1, the assumed condition $\alpha_n/k_n < \bar{\alpha}^n/\bar{k}^n$ implies $V_n - \pi_n^* = S_n + D_n > 0$. Thus, the interval $(F', F' + S_n + D_n)$ is non-empty. This completes the proof of Proposition 3. ■

Proof of Lemma 3

The firm's problem is again as in (A1). Given wholesale price $\tilde{w}_i$ and firm output q_i , rival i chooses $\tilde{q}_i$ to maximize its profit:

$$\text{Max}_{\tilde{q}_i} [a_i - \tilde{q}_i - \gamma_i q_i - \tilde{w}_i] \tilde{q}_i, \quad i = 1, \dots, n. \quad (\text{A17})$$

Jointly solving the first-order conditions of (A1) and (A17) yields:

$$q_i(w, \tilde{w}_i) = \frac{a_i[2 - \gamma_i] - 2w + \gamma_i \tilde{w}_i}{4 - \gamma_i^2}, \text{ and}$$

$$\tilde{q}_i(w, \tilde{w}_i) = \frac{a_i[2 - \gamma_i] - 2\tilde{w}_i + \gamma_i w}{4 - \gamma_i^2}, \quad i = 1, \dots, n. \quad (\text{A18})$$

The output reactions in (A18) serve as the induced demand function for the supplier, and it solves:

$$\text{Max}_{w, \tilde{w}_i, i=1, \dots, n} \sum_{i=1}^n [w q_i(w, \tilde{w}_i) + \tilde{w}_i \tilde{q}_i(w, \tilde{w}_i)]. \quad (\text{A19})$$

Solving the first-order conditions of (A19) yields w'^n and $\tilde{w}_i'^n$ in part (ii). Using these wholesale prices in (A18), the output decisions in part (i) equal $q_i(w'^n, \tilde{w}_i'^n)$ and $\tilde{q}_i(w'^n, \tilde{w}_i'^n)$. This completes the proof of Lemma 3. ■

Proof of Proposition 4

From Lemma 1, $q_i'^n = \frac{1}{2} \left[\frac{a_i}{2+\gamma_i} - \frac{\bar{\alpha}^n - a_i}{2} \right]$, $q_i'^{n-1} = \frac{1}{2} \left[\frac{a_i}{2+\gamma_i} - \frac{\bar{\alpha}^{n-1} - a_i}{2} \right]$, $w'^n = \frac{\bar{\alpha}^n}{2}$, and $w'^{n-1} = \frac{\bar{\alpha}^{n-1}}{2}$. Substituting these expressions in the analog to (A8), and making use of the identity $\bar{\alpha}^n = \frac{n-1}{n} \bar{\alpha}^{n-1} + \frac{1}{n} a_n$, algebra yields the expressions for S_n' and D_n' .

Using S_n' and D_n' :

$$V_n' - \pi_n'^* = \lambda' [\bar{\alpha}^n - a_n], \text{ where}$$

$$\lambda' = \left[\frac{1}{16(n-1)} \right] \left[8 \sum_{i=1}^{n-1} q_i'^n + a_n - \bar{\alpha}^n \right]. \quad (\text{A20})$$

Using $q_i'^n$ from Lemma 3(i), the λ' expression in (A20) can be equivalently written as follows:

$$\lambda' = \left[\frac{1}{16(n-1)} \right] \left[8 \sum_{i=1}^{n-1} \frac{a_i}{2(2+\gamma_i)} - \frac{8[(n-1)\bar{\alpha}^n - (n\bar{\alpha}^n - a_n)]}{4} + a_n - \bar{\alpha}^n \right]$$

$$= \left[\frac{1}{16(n-1)} \right] \left[4 \sum_{i=1}^{n-1} \frac{a_i}{2+\gamma_i} + \bar{\alpha}^n - a_n \right]. \quad (\text{A21})$$

If $a_n > \bar{\alpha}^n$, from (A20), $\lambda' > 0$. And, if $a_n < \bar{\alpha}^n$, from (A21), again $\lambda' > 0$. With $\lambda' > 0$, from (A20), the sign of $V_n' - \pi_n'^*$ is the same as the sign of $\bar{\alpha}^n - a_n$. This completes the proof of Proposition 4. ■

Proof of Lemma 4

Solving $p_i = a_i - q_i - \gamma_i \tilde{q}_i$ and $\tilde{p}_i = a_i - \tilde{q}_i - \gamma_i q_i$, the demand functions are $q_i(p_i, \tilde{p}_i) = \frac{a_i[1-\gamma_i]-p_i+\gamma_i\tilde{p}_i}{1-\gamma_i^2}$ and $\tilde{q}_i(p_i, \tilde{p}_i) = \frac{a_i[1-\gamma_i]-\tilde{p}_i+\gamma_i p_i}{1-\gamma_i^2}$. Given wholesale price w and rival price $\tilde{p}_i$, in the Bertrand game, the firm chooses p_i in segment i , $i = 1, \dots, n$, to maximize its total profit:

$$\text{Max}_{p_i, i=1, \dots, n} \sum_{i=1}^n [p_i - w] q_i(p_i, \tilde{p}_i). \quad (\text{A22})$$

Given retail price p_i , rival i chooses $\tilde{p}_i$ to maximize its profit:

$$\text{Max}_{\tilde{p}_i} \tilde{p}_i \tilde{q}_i(p_i, \tilde{p}_i), \quad i = 1, \dots, n. \quad (\text{A23})$$

Jointly solving the first-order conditions of (A22) and (A23) yields:

$$p_i(w) = \frac{a_i[2 - \gamma_i - \gamma_i^2] + 2w}{4 - \gamma_i^2}, \text{ and}$$

$$\tilde{p}_i(w) = \frac{a_i[2 - \gamma_i - \gamma_i^2] + \gamma_i w}{4 - \gamma_i^2}, \quad i = 1, \dots, n. \quad (\text{A24})$$

Using retail prices from (A24) in the demand functions $q_i(p_i, \tilde{p}_i)$ and $\tilde{q}_i(p_i, \tilde{p}_i)$ yields:

$$\hat{q}_i(w) = \frac{a_i[2 - \gamma_i - \gamma_i^2] - [2 - \gamma_i^2]w}{4 - 5\gamma_i^2 + \gamma_i^4} = \hat{\alpha}_i - w\hat{k}_i, \text{ and}$$

$$\hat{\tilde{q}}_i(w) = \frac{a_i[2 - \gamma_i - \gamma_i^2] + \gamma_i w}{4 - 5\gamma_i^2 + \gamma_i^4} = \hat{\alpha}_i + \frac{\gamma_i w \hat{k}_i}{2 - \gamma_i^2}, \quad i = 1, \dots, n. \quad (\text{A25})$$

(A25) serves as the induced demand function for the supplier, and it solves:

$$\text{Max}_w \sum_{i=1}^n w \hat{q}_i(w) \Leftrightarrow \text{Max}_w \sum_{i=1}^n w [\hat{\alpha}_i - w \hat{k}_i]. \quad (\text{A26})$$

Solving the first-order condition of (A26) yields $\hat{w}^n$ in part (ii). Using this wholesale price in (A25), the equilibrium outputs, noted in part (i), equal $\hat{q}_i(\hat{w}^n)$ and $\hat{\tilde{q}}_i(\hat{w}^n)$. This completes the proof of Lemma 4. ■

Proof of Proposition 5

From Lemma 4, $\hat{q}_i^n = \frac{1}{2} \left[2\hat{\alpha}_i - \frac{\hat{k}_i}{\hat{k}^n} \bar{\alpha}^n \right]$, $\hat{q}_i^{n-1} = \frac{1}{2} \left[2\hat{\alpha}_i - \frac{\hat{k}_i}{\hat{k}^{n-1}} \bar{\alpha}^{n-1} \right]$, $\hat{w}^n = \frac{\bar{\alpha}^n}{2\hat{k}^n}$, and $\hat{w}^{n-1} = \frac{\bar{\alpha}^{n-1}}{2\hat{k}^{n-1}}$. Substituting these in the analog to (A8), and making use of the identities $\bar{\alpha}^n = \frac{n-1}{n} \bar{\alpha}^{n-1} + \frac{1}{n} \alpha_n$ and $\hat{k}^n = \frac{n-1}{n} \hat{k}^{n-1} + \frac{1}{n} k_n$, algebra yields the expressions for $\hat{S}_n$ and $\hat{D}_n$.

Using $\hat{S}_n$ and $\hat{D}_n$:

$$\hat{V}_n - \hat{\pi}_n^* = \hat{\lambda} \left[\frac{\bar{\alpha}^n}{\bar{k}^n} - \frac{\hat{\alpha}_n}{\hat{k}_n} \right], \text{ where}$$

$$\hat{\lambda} = \left[\frac{\hat{k}_n}{2(n-1)\bar{k}^{n-1}} \right] \left[2 \sum_{i=1}^{n-1} \frac{(2-\gamma_i^2)\hat{q}_i^n}{4-\gamma_i^2} - \left(\frac{\hat{k}_n \bar{\alpha}^n}{\bar{k}^n} - \hat{\alpha}_n \right) \left(\frac{\sum_{i=1}^{n-1} (1-\gamma_i^2)\hat{k}_i^2}{2(n-1)\bar{k}^{n-1}} \right) \right]. \quad (\text{A27})$$

The sign of $\hat{V}_n - \hat{\pi}_n^*$ is the same as the sign of $\frac{\bar{\alpha}^n}{\bar{k}^n} - \frac{\hat{\alpha}_n}{\hat{k}_n}$ if $\hat{\lambda} > 0$. Note that $\hat{\lambda} > 0$ if and only if the term $\hat{T} = 2 \sum_{i=1}^{n-1} \frac{(2-\gamma_i^2)\hat{q}_i^n}{4-\gamma_i^2} - \left(\frac{\hat{k}_n \bar{\alpha}^n}{\bar{k}^n} - \hat{\alpha}_n \right) \left(\frac{\sum_{i=1}^{n-1} (1-\gamma_i^2)\hat{k}_i^2}{2(n-1)\bar{k}^{n-1}} \right) > 0$.

Because $\hat{q}_i^n > 0$, it follows that if $\hat{\alpha}_n > \frac{\hat{k}_n \bar{\alpha}^n}{\bar{k}^n}$, $\hat{T} > 0$. Thus the proof is complete if it can be shown that $\hat{T} > 0$ even when $\hat{\alpha}_n < \frac{\hat{k}_n \bar{\alpha}^n}{\bar{k}^n}$. Given $\hat{\alpha}_n < \frac{\hat{k}_n \bar{\alpha}^n}{\bar{k}^n}$ and $1/3 \leq \frac{2-\gamma_i^2}{4-\gamma_i^2} \leq 1/2$ (with the limiting values corresponding to $\gamma_i = 0$ and $\gamma_i = 1$), a lower bound on $\hat{T}$ is derived below:

$$\begin{aligned} \hat{T} &= 2 \sum_{i=1}^{n-1} \frac{(2-\gamma_i^2)\hat{q}_i^n}{4-\gamma_i^2} - \left(\frac{\hat{k}_n \bar{\alpha}^n}{\bar{k}^n} - \hat{\alpha}_n \right) \left(\frac{\sum_{i=1}^{n-1} [(2-\gamma_i^2)/(4-\gamma_i^2)]\hat{k}_i^2}{2(n-1)\bar{k}^{n-1}} \right) > \frac{2}{3} \sum_{i=1}^{n-1} \hat{q}_i^n - \left(\frac{\hat{k}_n \bar{\alpha}^n}{\bar{k}^n} - \hat{\alpha}_n \right) \\ &\times \left(\frac{\sum_{i=1}^{n-1} \hat{k}_i^2}{4(n-1)\bar{k}^{n-1}} \right) = \frac{2}{3} \sum_{i=1}^{n-1} \hat{q}_i^n - \frac{1}{4} \left(\frac{\hat{k}_n \bar{\alpha}^n}{\bar{k}^n} - \hat{\alpha}_n \right) = \frac{2}{3} \left(\frac{n\bar{\alpha}^n}{2} - \hat{\alpha}_n + \frac{\hat{k}_n \bar{\alpha}^n}{2\bar{k}^n} \right) - \frac{1}{4} \left(\frac{\hat{k}_n \bar{\alpha}^n}{\bar{k}^n} - \hat{\alpha}_n \right) \\ &= \frac{1}{3} (n\bar{\alpha}^n - \hat{\alpha}_n) + \frac{1}{12} \left(\frac{\hat{k}_n \bar{\alpha}^n}{\bar{k}^n} - \hat{\alpha}_n \right) = \frac{1}{3} \left(\sum_{i=1}^{n-1} \hat{\alpha}_i \right) + \frac{1}{12} \left(\frac{\hat{k}_n \bar{\alpha}^n}{\bar{k}^n} - \hat{\alpha}_n \right). \quad (\text{A28}) \end{aligned}$$

From (A28), for $\hat{\alpha}_n < \frac{\hat{k}_n \bar{\alpha}^n}{\bar{k}^n}$, $\hat{T} > 0$ and, hence, $\hat{\lambda} > 0$. This completes the proof of Proposition 5. ■

REFERENCES

- Arya, A., and B. Mittendorf. 2006. Benefits of channel discord in the sale of durable goods. *Marketing Science* 25 (1): 91–96.
- , ———, and D. Sappington. 2008. The make-or-buy decision in the presence of a rival: Strategic outsourcing to a common supplier. *Management Science* 54 (10): 1747–1758.
- Baiman, S., P. Fischer, and M. Rajan. 2000. Information, contracting, and quality costs. *Management Science* 46 (6): 776–789.
- , ———, and ———. 2001. Performance measurement and design in supply chains. *Management Science* 47 (1): 173–188.
- , and M. Rajan. 2002. The role of information and opportunism in the choice of buyer-supplier relationships. *Journal of Accounting Research* 40 (2): 247–278.

- Balachandran, K., and S. Radhakrishnan. 2005. Quality implications of warranties in a supply chain. *Management Science* 51 (8): 1266–1277.
- Chevalier, J. 2004. What do we know about cross-subsidization? Evidence from merging firms. *Advances in Economic Analysis & Policy* 4.
- Chozick, A. 2007. The samurai sell: Lexus dealers bow to move swank cars; in Japan, Toyota revives ancient hospitality rites, but some buyers balk. *Wall Street Journal* (July 9): A1.
- Cooper, R., and R. Kaplan. 1988. Measure costs right: Make the right decisions. *Harvard Business Review* 66 (September–October): 96–103.
- Darrough, M. 1993. Disclosure policy and competition: Cournot vs. Bertrand. *The Accounting Review* 68 (3): 534–561.
- Dye, R. 2001. An evaluation of “essays on disclosure” and the disclosure literature in accounting. *Journal of Accounting and Economics* 32 (1–3): 181–235.
- Gertner, R., E. Powers, and D. Scharfstein. 2002. Learning about internal capital markets from corporate spin-offs. *The Journal of Finance* 57 (6): 2479–2506.
- Goebel, D., G. Marshall, and W. Locander. 1998. Activity-based costing accounting for a market orientation. *Industrial Marketing Management* 27 (6): 497–510.
- Hogan, J., D. Lehmann, M. Merino, R. Srivastava, J. Thomas, and P. Verhoef. 2002. Linking customer assets to financial performance. *Journal of Service Research* 5 (1): 26–38.
- Hwang, I., S. Radhakrishnan, and L. Su. 2006. Vendor certification and appraisal: Implications for supplier quality. *Management Science* 52 (10): 1472–1482.
- Katz, M. 1989. Vertical contractual relations. In *Handbook of Industrial Organization*, edited by R. Schmalensee, and R. Willig., Vol. 1. 655–721 London, U.K.: Elsevier.
- Kuchta, D., and M. Troska. 2007. Activity-based costing and customer profitability. *Cost Management* 21 (3): 18–25.
- Lariviere, M. 2008. Supply chains. In *The New Palgrave Dictionary of Economics*, edited by L. Blume, and S. Durlauf. 2nd edition. New York, NY: Palgrave Macmillan.
- Rajan, R., H. Servaes, and L. Zingales. 2000. The cost of diversity: The diversification discount and inefficient investment. *The Journal of Finance* 55 (1): 35–80.
- Saouma, R. 2008. Optimal second stage outsourcing. *Management Science* 54 (6): 1147–1159.
- Searcy, D. 2004. Using activity-based costing to assess channel/customer profitability. *Management Accounting Quarterly* 5 (Winter): 51–60.
- Shin, H., and R. Stulz. 1998. Are internal capital markets efficient? *The Quarterly Journal of Economics* 113 (2): 531–552.
- Sopariwala, P. 2005. Evaluating the “keep-or-drop” decision. *Journal of Corporate Accounting & Finance* 16 (3): 47–61.
- Spengler, J. 1950. Vertical integration and antitrust policy. *The Journal of Political Economy* 58 (4): 347–352.
- Su, X., and F. Zhang. 2008. Strategic customer behavior, commitment, and supply chain performance. *Management Science* 54 (10): 1759–1773.
- Venkatesan, R., and V. Kumar. 2004. A customer lifetime value framework for customer selection and resource allocation strategy. *Journal of Marketing* 68 (4): 106–125.
- Verrecchia, R. 2001. Essays on disclosure. *Journal of Accounting and Economics* 32 (1–3): 97–180.
- Zimmerman, J. 2003. *Accounting for Decision Making and Control*. New York, NY: McGraw-Hill Irwin.

Detection and Severity Classifications of Sarbanes-Oxley Section 404 Internal Control Deficiencies

*Jean C. Bedard
Lynford Graham
Bentley University*

ABSTRACT: We examine detection and severity classification of Internal control deficiencies (ICD) identified under Section 404 of the Sarbanes-Oxley Act of 2002. While the cost/benefit balance of auditor testing of internal controls is highly controversial, prior research has not examined auditor versus client detection of ICD, nor has it examined factors auditors consider in judging ICD severity. We find that auditors detect about three-fourths of unremediated ICD, usually through control testing. This finding contrasts with extant research inferring control deficiency detection effectiveness from publicly available data, underscoring the value of Section 404 auditor testing in improving financial reporting quality. Auditors judge greater severity when a misstatement has already occurred. In the absence of a misstatement, severity is contingent on client and ICD characteristics, implying a more complex and nuanced judgment process without objective evidence of control failure. We also find that clients often underestimate ICD severity, but this tendency is lower among well-controlled companies with a well-designed Section 404 process.

Keywords: *Internal controls; Sarbanes-Oxley Section 404; risk assessment; materiality.*

I. INTRODUCTION

This study investigates detection and severity classification of internal control deficiencies (ICD) under Section 404 of the Sarbanes-Oxley Act (SOX; U.S. Congress 2002). The purpose of Section 404 is to improve the reliability of information provided by public companies to the financial markets (COSO 2006; PCAOB 2004) by requiring company manage-

We thank the participating firms for their support of this research. This paper has benefited from comments of participants of the 2008 International Symposium on Auditing Research (Pasadena), the 2008 Conference of the Accounting and Finance Association of Australia and New Zealand (Sydney), the 2008 American Accounting Association Annual Meeting (Anaheim), and accounting research workshops at Bentley University, the College of William and Mary, and Northeastern University. We also appreciate helpful comments from Mohammad Abdolmohammadi, Carl Brewer, Christine Earley, Rani Hoitash, Udi Hoitash, Jim Irving, Karla Johnstone, Bill Kinney, Ganesh Krishnamoorthy, Mario Maletta, Tom Omer, Margie Platt, Tim Rupert, Roger Simnett, Stephen Taylor, Stuart Taylor, Jay Thibodeau, Kim Westermann, Arnie Wright, two anonymous reviewers, and Steve Kachelmeier (senior editor).

Editor's note: Accepted by Steven Kachelmeier.

*Submitted: January 2009
Accepted: October 2010
Published Online: April 2011*

ment and auditors to test internal control over financial reporting (ICFR) and to disclose severe control flaws that are not remediated as of the balance sheet date. Disclosure of material weakness(es) (MW) involves both detection and classification (Leone 2007), but in order to isolate these processes for study, information is needed on the base of ICD from which MW are derived.

Using proprietary data on detected ICD provided by several large audit firms, we address two research questions. First, what is the relative contribution of clients and auditors to ICD detection? This study is unique in its ability to examine, within the same time period, the extent of client (Section 404(a)) versus auditor (Section 404(b)) detection of ICD when both parties document and test controls. Within this analysis we examine factors contributing to client and auditor detection. One key feature of our study is that we provide direct evidence on the “yield” of detection methods used by auditors. This is an important issue at the heart of the debate on the value of auditor involvement in assessing and testing internal controls. Non-accelerated filers have recently been exempted by Congress from auditor attestation of controls effectiveness (Section 404(b)) as part of the financial reform legislation (U.S. Congress 2010). Emerging research (Kinney and Shepardson 2010; hereafter, KS) addresses the impact of auditor testing of controls using publicly available data. KS infer that because Audit Analytics (AA) data show that Section 404 reports for many companies with MW mention “material and/or numerous year-end adjustments,” substantive testing might be sufficient to detect at least one MW for such companies, resulting in public reporting of ineffective controls. However, due to limitations of publicly available data, KS are not able to track misstatements to individual MW and, thus, their conclusion relates to company-level reporting of effective or ineffective controls. In contrast, we analyze deficiencies at the level of the individual control, directly examining detection method based on reports by engagement teams (as opposed to inferring possible detection method from annual public reports). Further, our data include not only MW, but also other detected ICD judged by auditors to be less severe.

Second, what factors are associated with the auditor’s severity classifications of detected ICD? These classifications determine which control flaws are publicly reported (MW), and which are communicated to management and the audit committee (significant deficiency/deficiencies; SD).¹ In AS No. 2 (PCAOB 2004), SD are those deficiencies with a “more than remote” likelihood of failing to prevent/detect misstatements of “more than inconsequential” amount.² For MW, the amount of potential misstatement must be “material.” These criteria for judging misstatement likelihood and materiality are likely complex to apply in practice due to difficulty in interpreting probability phrases, as recognized by prior research (Amer et al. 1994; Asare and Wright 2008). In addition, the wording of classification criteria is ambiguous, in implying that a deficiency can be “significant” (generally defined as important or consequential), but not be “material” (capable of influencing the judgments of financial report users). We study auditor judgments in the complex, unstructured task of severity classification, in this context of imprecise definitions and guidance.

We investigate these issues in a sample of audits of smaller accelerated filer companies, in which 3,990 ICD were detected and 2,942 remained unremediated at year-end.³ Regarding our

¹ Both classifications are consequential. Prior research shows that MW disclosure under Sections 404 or 302 is associated with adverse consequences to companies (e.g., Krishnan et al. 2008; Hammesley et al. 2008), and to company officers (Li et al. 2010). Because SD are not known unless voluntarily disclosed, research has not studied adverse consequences to company officers from SD detection. However, SD are also important because they may aggregate to a MW, or evolve into MW if not remediated in a timely manner (AS No. 2, ¶140).

² AS No. 2, effective during our study period, was replaced by AS No. 5 (PCAOB 2007), which also requires that management’s assertion and the auditor’s opinion on ICFR effectiveness are supported by testing of controls. At inception in 2004, Section 404 applied to “accelerated filers,” i.e., companies with \$75 million of common equity float, among other criteria. Also during our study period, the SEC suggested that company managements look to PCAOB guidance for ICD classification. Specific SEC guidance for management was issued after our study period.

³ Severity classifications can only be studied for unremediated ICD. In order to examine detection and classification in the same sample, we restrict the detection analysis to unremediated ICD.

first research question, we find that clients detect fewer ICD than auditors, and are less likely to detect severe and pervasive ICD. Thus, many control flaws most likely to affect financial reporting would not be found in a client-driven process such as Section 302. Our results also show that use of a large accounting firm consultant for Section 404(a) work is associated with improved client detection. We extend this analysis by investigating how auditors detect ICD. We find that control tests provide initial evidence on a large proportion of ICD, including most MW and entity-level problems viewed as more serious by financial report users (Hammersley et al. 2008). This finding affirms auditors' Section 404 control testing as an important source of detecting control deficiencies, and implies that the recent exemption of Section 404(b) for smaller U.S. public companies could result in failure to fully realize potential improvements in financial reporting quality in that sector of the market.

Our second research question considers ICD severity classification. We first investigate the influence of a preliminary client classification (Earley et al. 2008). We find that clients tend to classify ICD as less severe, but auditors frequently override those classifications. Modeling ICD severity classifications, we find higher severity associated with: (1) greater knowledge and independence in the client's Section 404(a) process; (2) more objective evidence (e.g., an existing misstatement); (3) control flaws other than documentation problems (e.g., inappropriate design); (4) certain types of entity-level ICD (e.g., Control Environment); and (5) certain types of account-specific ICD (revenue and tax), consistent with the regulatory climate of the period.

Because prior research underscores the importance of objective evidence when standards are imprecise (e.g., Nelson et al. 2002), we further explore the role of an associated misstatements in ICD classification. An associated misstatement provides objective evidence that an ICD may fail again in the near future, but how do auditors judge severity when a misstatement has not already occurred? In this situation, we find that an independent client-reporting process, a longer auditor testing period, substantive test detection, and the number of MW identified become more important in classifying a given ICD as severe. Simply put, when objective evidence from a misstatement is not available, auditors use other evidence to support the classification decision.

In summary, our results support the value of auditor involvement at two stages of the ICFR assessment process (detection and classification), and contribute to understanding of factors associated with client and auditor performance in both stages. The following section provides background and contribution of this study, and develops our research questions and specific expectations. Section III discusses the study's methods, and Section IV contains results of models and other analyses. Section V presents the study's conclusions and limitations.

II. BACKGROUND, PRIOR RESEARCH, AND DEVELOPMENT OF RESEARCH QUESTIONS

Internal Control Regulation in the U.S. and Prior Related Research

SOX was implemented in response to numerous large-scale frauds and financial restatements among public companies. Among its key provisions is Section 404, which is based on the principle that improving ICFR should increase information reliability and restore investor confidence. This provision requires large public companies and their auditors to attest to ICFR effectiveness. Auditing standards prior to SOX included an assessment of internal controls as an element in obtaining the required understanding of the audited company and its environment.⁴ Under those stan-

⁴ An assessment of the control environment as a basis for risk assessment and planning is required in SAS No. 55 (AICPA 1988). SAS No. 78 (AICPA 1997) builds on that standard by incorporating the five components of the COSO Framework. Under AU 325, auditors were required to disclose observed "reportable conditions" (including material weaknesses in controls) to management and the audit committee or equivalent, but such communications were not made public. In 2003, the PCAOB adopted the AICPA Standards issued to date as interim standards.

dards, auditors had the option of relying on controls that were tested and found effective, or primarily using substantive procedures to support the financial statement audit opinion. Under prior standards, few auditors were familiar with documenting and testing of controls at the level expected by Section 404, using the COSO framework, and with the level of assurance required to issue a report on control effectiveness (see also KS; Curtis et al. 2009).

Prior research uses publicly available annual report data to distinguish characteristics of companies disclosing ineffective controls (i.e., at least one MW) under Section 404, or quarterly management reports under Section 302 (Ashbaugh-Skaife et al. 2007; Doyle et al. 2007; Hoitash et al. 2009). However, archival research does not address the full extent of detected control flaws, how those problems are detected, or how auditors determine which problems are disclosed. For example, KS use publicly available data on misstatements mentioned in Section 404 of the annual reports to infer that most companies would have disclosed ineffective controls (i.e., at least one MW detected) if a substantive approach were used. Because studies using publicly available data cannot directly examine detection method, however, research with access to control-level data is important. Further, studies using public reports have an incomplete data set from which to draw conclusions on detection method or the auditor's role, as only disclosure of unremediated MW is required. Evidence on auditors' severity classification decisions from behavioral experiments highlights the judgmental nature of these decisions. Earley et al. (2008) find that clients' preliminary severity classifications influence auditor classifications, which is especially problematic when clients propose low severity. Also, Wolfe et al. (2009) show that management persuasion tactics can affect auditors' classification of information technology controls.

In general, prior research implies that the component tasks of ICD severity classification (judging misstatement likelihood and materiality) are unstructured, complex, and difficult (e.g., Messier et al. 2005; Allen et al. 2006). One source of difficulty is interpretation of phrases such as "more than remote" and "material," derived from guidance on financial statement materiality (SAB No. 99; SEC 1999a) and accounting for contingencies (SFAS No. 5; FASB 1975). Another is inconsistency in definitions. In defining MW, the PCAOB follows the FASB's definition of "material" (SFAC No. 2, ¶132) as something that would change the judgment of a reasonable person relying on the information. Yet "significant" in the title of SD generally means something that is important, consequential, and large. Thus, auditing standards imply that an ICD can be important (thus a SD) but not affect users' decisions. Both issues lead to "ambiguity in the definitions" (Leone 2007, 227) of SD and MW, allowing latitude for company management in ICD classification. This ambiguity makes it difficult for auditors to judge the "right" ICD classification, and to justify that decision to client personnel whose incentive is not to report MW or SD. But to adhere to professional standards in a climate of liability concerns and PCAOB inspections, the auditor's incentive is to detect existing flaws in ICFR and bring them to the attention of client management, the audit committee, and the public if necessary. This study provides the first archival evidence of ICD detection and severity classification. In the following sections, we develop our research expectations from prior literature, auditing standards, and the economic context of the early years of Section 404 implementation.

Research Expectations

ICD Detection

We first focus on ICD detection, expecting that detection quality increases with expertise and independence of the parties performing the work. We expect that auditors will detect more ICD than clients, as they are independent from clients and have more knowledge of internal controls derived from education, firm training, and experience. We expect more detection among companies using Section 404 consultants, who are hired for their expertise and, in addition, have more independence from management than client personnel (although less than the external auditor). We

expect a greater impact when the consultant is from a large auditing firm due to significant training and guidance received in those firms. The relative knowledge of consultant personnel of other entities is less clear, as smaller firms likely did not institute the degree of training used by larger firms. In terms of relative knowledge/independence, we expect that companies with internal audit functions will outperform companies with no internal audit function.

In addition to these indicators of *who* is performing the client's control documentation and testing, we also consider the level of independence built into the client's process design. If company management is not involved in the reporting chain, then there is less opportunity for management influence. Thus, we expect a positive association of Section 404(a) process independence with client detection of ICD. We also consider the impact of integration of information technology (IT) in the Section 404(a) process, but do not predict a sign. While clients utilizing more advanced technology may detect more control flaws, those employing more IT may also have better underlying systems (i.e., fewer manual controls). Finally, we anticipate that a longer Section 404(a) testing period will result in better ICD detection.

We also investigate the nature of ICD detected by clients, an issue with important implications for understanding which control flaws are likely to be missed under the current regulatory structure for non-accelerated filers, which does not require auditor involvement. We expect that clients will be less likely to discover more severe problems, due to relative lack of knowledge about controls and the inherent objectivity challenge of "self-assessment." The issue of independence also arises in connection with client detection of entity-level ICD. Because many entity-level problems occur in centralized functions that directly involve top management, we expect that company personnel will be less likely to recognize flaws in controls at that level, relative to key Control Activities in areas such as sales and purchases.

ICD Severity Classification

Client characteristics and Section 404 processes. Our second summary research question concerns factors affecting auditors' severity classification decisions. We first consider client characteristics. Specifically, we expect higher ICD severity classifications in companies with lower overall quality of internal control.⁵ In companies with good controls, an individual ICD is an isolated incident. However, in a weak-control context, auditors may consider a given ICD to be more severe, due to the potential for interaction with other control flaws and the potential for less effective compensating controls. We derive evidence on the overall controls context from the auditor's control reliance decision. We also consider the number of MW detected as an indicator of pervasive and serious problems in ICFR specifically (see also Hoitash et al. 2008; Hermanson and Ye 2009).

Our second severity expectation is that classifications will vary based on strength of the client's Section 404(a) processes. Consultants' greater expertise and independence could lead to better recognition of actual ICD severity, but these qualities could also result in remediation of more severe problems prior to year-end. Thus, we do not predict a directional effect for large firm or other Section 404 consultants. In addition, we expect that client Section 404(a) process independence will be positively associated with severity, as results of Section 404(a) activity will not be filtered through company management, which has the incentive not to report SD or MW.

We also consider the role of internal audit, proposing a nondirectional expectation. On the one hand, internal audit personnel may be more adept at recognizing ICD severity than other client personnel, if the internal audit unit has been involved in controls design. However, companies with

⁵ Prior research (Ashbaugh-Skaife et al. 2007) tests the association of control context with MW disclosure indirectly, e.g., using size and financial position to proxy for investment in controls. We build on that study by isolating severity classification decisions and using more precise measures.

an internal audit could have less severe control flaws (Felix et al. 2001; Lin et al. 2011). We also consider the extent to which IT is integrated into Section 404(a) activity. Better use of IT could result in detection of more severe problems, as documentation and testing of controls will be more complete and accurate than a manual process. However, companies with greater use of IT for Section 404(a) likely also have better use of IT throughout their operations and, hence, may have better controls. Thus, we also do not predict a sign on this variable.

Auditor Section 404 processes. The auditor's Section 404(b) processes could also affect severity classification. First, we expect that severity classifications will be higher for auditor-detected ICD. Given their greater knowledge and objectivity, auditors should detect more serious control flaws than clients, and should be more confident of the severity implications of those they detect. The second factor in this category is the length of the auditor's testing process. During our sample period, anecdotal evidence from audit partners noted concern that clients' delay in completing their work could affect the quality of the auditor's process that follows. Further, due to constraints in the audit labor market during our study period, firms with heavy concentrations of accelerated filer clients struggled to hire and retain sufficient numbers of qualified employees to perform Section 404 procedures. Thus, an earlier start for auditor control testing should improve engagement teams' ability to fully consider and properly classify ICD.

Third, we consider the auditor's source of evidence. Specifically, we expect higher severity classifications associated with detection of ICD through substantive tests, and with an existing misstatement associated with the flawed control. Substantive tests are designed to detect account discrepancies that are the basis for proposed audit adjustments. If a monetary discrepancy has been detected, then evidence of ICD severity is stronger with respect to the likelihood criterion. Also, an actual misstatement amount creates a "floor" for the quantitative effect of the control deficiency, even though the potential severity could be greater than the amount observed. The presence of a misstatement is important, given the widespread use of quantitative criteria for determining materiality (e.g., Ng and Tan 2003; Messier et al. 2005). The association of misstatements with severity of control flaws is important to the argument proposed by KS that MW for many companies can be detected by tracing back from misstatements. If detection is through control tests and no associated misstatement has been detected, then the auditor's judgment in the Section 404 context becomes more complex—s/he must consider whether financial statement users' decisions could be affected by a *possible future* misstatement arising from a recognized ICD, without benefit of an actual current misstatement available as an example.

Prior research (Kennedy et al. 1997; Nelson et al. 2002; DeZoort et al. 2003) supports a positive association of misstatements with severity classification, by finding that objective evidence is important in justifying auditor decisions in unstructured tasks. These studies suggest that tangible evidence of control failure will improve auditors' ability to defend ICD classification decisions to management and audit committees, as well as in subsequent internal firm and PCAOB inspections. But how do auditors make severity classification decisions in the absence of a misstatement? We expect that in this case, stronger client and auditor processes will help bolster the auditor's position in discussions with management and the audit committee over ICD classification. Thus, process features such as an independent reporting chain and a longer auditor testing period could be more significant when a misstatement has not already occurred.

ICD type. We next consider whether severity classifications vary according to the nature of the ICD.⁶ AS No. 2, ¶133 notes that the auditor should consider the nature of the affected financial

⁶ Motivating study of differential classification according to nature of the control flaw, professional standards on Section 404 require specific consideration of accounts and assertions (AS No. 2, paras. 68–69). While past studies find little difference in control risk assessments across accounts/assertions (e.g., Elder and Allen 2003), current standards suggest that we could find systematic differences in severity classifications across types of ICD.

statement accounts in assessing the likelihood of misstatements that may result from an inoperative control. AS No. 2, ¶50–53 focus on entity-level controls with pervasive effects, including IT general controls (ITGC), controls over the financial reporting close process, and COSO categories of Control Environment, Monitoring, and company Risk Assessment. While some studies address the importance of Control Environment in control risk assessments (e.g., Cohen and Hanno 2000; Cohen et al. 2002), prior research has not directly examined how entity-level status affects auditors' severity classification in the Section 404 context.

Among account-specific ICD, we expect revenue accounts to be considered more severe. The SEC's strong focus on revenue recognition in the period immediately preceding the enactment of SOX (e.g., SAB No. 101; SEC 1999b) heightened auditors' awareness of the importance of controls over revenue recognition. We also predict a positive association of tax ICD with severity. In response to concerns about auditor independence, the SEC prohibited auditors from performing certain accounting services following passage of SOX, including preparation of the tax accrual. The transition away from auditor assistance to companies in tax accrual accounting to meet SEC/PCAOB independence requirements suggests higher likelihood of a material misstatement in tax accounts during our study period.

Professional guidance and some prior research on the second component judgment, materiality, also suggests differential classification according to the nature of the account. The definition of materiality in professional auditing standards is based on the auditor's judgment that an item would make a difference in users' decisions. AS No. 2, ¶23 notes that the auditor should consider both quantitative and qualitative elements, as in the financial statement materiality judgment (SAB No. 99; SEC 1999a). Qualitative factors noted in the standard include the nature of accounts and assertions, and future misstatement consequences. Prior research mainly considers materiality judgments in audit planning and audit adjustments. Summarizing this literature, Messier et al. (2005) conclude that during planning, the primary criterion for materiality in practice is income-based, although other criteria are considered. As noted above, judgments of materiality are likely more difficult when an existing misstatement is not available to provide a gauge for the likely amount of future misstatements resulting from the ICD.

In sum, this research builds on prior studies of publicly disclosed Section 404 MW by investigating detection and severity classifications of ICD discovered by companies and auditors. In the following section, we discuss our procedures and methods of testing expectations.

III. METHOD

Sample and Data Collection Procedures

We obtained data from several large audit firms, under confidentiality agreements that restrict us from knowing client identities and reporting separate results by firm or firm size. To guide sample selection, we asked that each firm randomly select from 2004–2005 engagements of smaller accelerated filers (with revenues of about \$1 billion or less) in non-regulated industries. The restriction on client size is to improve our ability to generalize to the more than 10,000 smaller U.S. public companies not subject to Section 404(b) and to listed companies in other countries. Contact personnel from participating firms helped develop a spreadsheet to be completed by engagement teams, containing both company-level and control-level information.

Variable Definitions and Coding

Table 1, Panel A describes company-level variables representing auditors' and clients' Section 404 processes and outcomes, and Panel B describes control-level variables. We measure control quality using *IC_RELIANCE*, the auditor's level of reliance on internal controls. This is an indi-

TABLE 1
Client Controls Quality and Detection Process Measures
Panel A: Company-Level Client Control and Detection Process Variables: Mean (Std. Dev.) or Percent = 1

	All Engagements (n = 76)	Effective Controls (n = 61)	Ineffective Controls (n = 15)
<i>IC_RELIANCE</i> = 1 if current-year reliance on internal controls is "strong" or "maximum," and 0 otherwise.	48.7%	54.1%*	26.7%
<i>CONSULTANT_BIG6</i> = percent of clients using a large audit firm as consultant for Section 404(a) activity.	13.2%	13.1%	13.3%
<i>CONSULTANT_OTHER</i> = percent of clients using a consultant other than a large audit firm.	63.2%	65.6%	53.3%
<i>REPORT_INDEPENDENT</i> = results of Section 404(a) activity are reported to the audit committee or auditor, independent of company management.	34.2%	32.8%	40.0%
<i>INTERNAL_AUDIT</i> = 1 if client has an internal audit function, and 0 otherwise.	46.1%	44.3%	53.3%
<i>IT_INTEGRATION</i> = auditor's judgment of the effectiveness of the client's integration of information technology into its 404(a) process (5 = highly effective to 1 = highly ineffective).	3.30 (0.88)	3.39 (0.88)	2.93 (0.80)**
<i>MONTHS_PRIOR_CLIENT</i> = months before year-end that client personnel began Section 404(a) testing.	5.82 (2.23)	5.91 (2.28)	5.47 (2.03)
<i>MONTHS_PRIOR_AUDITOR</i> = months before year-end auditor began Section 404(b) testing.	4.21 (1.76)	4.13 (1.70)	4.53 (2.03)
<i>NUM_MW</i> = number of material weaknesses.	2.01 (9.00)	0.00 (0.00)	10.20 (18.55)***
<i>NUM_SD</i> = number of significant deficiencies.	6.07 (14.06)	3.07 (4.88)	18.27 (28.61)***
<i>NUM_DEF</i> = number of deficiencies.	30.63 (41.76)	27.82 (38.35)	42.07 (53.51)

(continued on next page)

TABLE B: CONTROL-LEVEL VARIABLES (continued)

	Overall (n = 3,990)	Deficiencies (n = 2,328)	Significant Deficiencies (n = 461)	Material Weaknesses (n = 153)	Remediated (n = 1,027)
<i>DETECTED_AUDITOR</i> = identified by the auditor (or auditor and client jointly).	72.1%	74.7%	80.9%	84.3%	60.6%
<i>DETECTED_SUBS</i> = detected by substantive tests.	9.8%	9.2%	17.4%	26.1%	5.1%
<i>DETECTED_CONTROL</i> = detected by control tests.	63.5%	68.5%	67.5%	58.8%	56.1%
<i>MISSTATEMENT_ASSOCIATED</i> = ICD has failed to prevent/detect a misstatement, resulting in a current audit adjustment or restatement of prior periods.	11.2%	9.1%	20.0%	27.5%	9.6%
<i>ENTITY</i> = ICD at the entity level (COSO Control Environment, Monitoring, Information/Communications and Risk Assessment; General Ledger Controls, and Information Technology General Controls).	55.7%	55.6%	62.0%	46.1%	55.2%

*, **, *** Indicate ≤ 0.10 , ≤ 0.05 , and ≤ 0.01 , respectively, with one-tailed probability levels for t-tests.

This table presents descriptive statistics on company-level and control-level variables.

Panel A presents company-level variables on client controls quality and detection processes of clients and auditors, with differences between engagements with effective/ineffective controls tested using independent sample t-tests or Pearson χ^2 tests (with d.f. = 1).

Panel B presents percentages of control-level variables, by severity category.

cator variable equaling 1 if “strong” or “maximum,” and 0 otherwise.⁷ *CONSULTANT_BIG6* equals 1 if the client outsourced all or part of its Section 404(a) activity to one of the largest six accounting firms, and 0 otherwise. *CONSULTANT_OTHER* equals 1 if the client outsourced its Section 404(a) activity to a smaller accounting firm or other provider. *REPORT_INDEPENDENT* equals 1 if Section 404(a) results are reported to the audit committee and auditor independent of client management, and 0 otherwise. *INTERNAL_AUDIT* equals 1 if the company has an internal audit function, and 0 otherwise. We capture the auditor’s assessment of the company’s effectiveness of integrating IT using a five-point scale (*IT_INTEGRATION*). *MONTHS_PRIOR_CLIENT* and *MONTHS_PRIOR_AUDITOR* measure the number of months before the balance sheet date that clients (auditors) began tests of controls. Table 1, Panel A also describes measures of the number of ICD at each severity level per engagement: *NUM_MW*, *NUM_SD*, and *NUM_DEF* for MW, SD, and deficiencies, respectively.

Control-level variables are described in Table 1, Panel B. Column headings indicate whether the ICD was remediated by year-end, and if not, the auditor’s final severity classification. If the ICD was detected by the client, then the instrument asks for the client’s preliminary severity classification, if known. *DETECTED_AUDITOR* equals 1 if detected by the auditor (or auditor and client jointly).⁸ We also asked firms to indicate the phase of the audit in which the ICD was detected (planning, internal control assessment/testing, substantive testing, or review). *DETECTED_SUBS* equals 1 if the ICD was detected by substantive tests, and 0 otherwise; *DETECTED_CONTROL* equals 1 if detected by substantive tests, and 0 otherwise. If the ICD was associated with a misstatement, then *MISSTATEMENT* equals 1, and 0 otherwise.⁹ We developed additional coding rules from professional sources, verified by contact personnel at the firms providing data. We classified each ICD by the COSO components: Control Environment, Monitoring, Risk Assessment, Information/Communications, and Control Activities. We use AS No. 2 to determine which ICD are entity-level; i.e., those in the first four COSO categories (*ENTITY_CONTROLENVIR*, *ENTITY_MONITORING*, *ENTITY_INFOCOM*, and *ENTITY_RISKASSESS*), IT general controls (*ENTITY_ITGC*), and general ledger controls (*ENTITY_GL*). The indicator *ENTITY* equals 1 if the ICD is in any entity category, 0 otherwise.¹⁰

When applicable, ICD are classified by financial statement accounts/cycle, based on the COSO (2006) small business framework and guidance from some participating firms. These are: *ACC_PURCHASES*, *ACC_REVENUES*, *ACC_COMPENSATION*, *ACC_FIXEDINTANG* (fixed and intangible assets), *ACC_TAXES*, *ACC_LIABILITIES* (including leases, warranty and legal liabilities), and *ACC_TREASURY* (including investments and bank account management). We also classified each ICD by the nature of the control weakness, with categories derived from COSO and

⁷ Control reliance was assessed using the following scale: 5 = maximum reliance; 4 = significant reliance; 3 = some reliance; 2 = no reliance; 1 = not applicable. Information on current-year reliance is missing for ten engagements. To enable use of this variable, we construct an indicator equaling 1 for maximum and significant reliance, and 0 otherwise. Measurement error from collapsing categories should bias against finding an effect of control reliance.

⁸ While auditing standards indicate that auditor testing should follow client testing within a given area, engagement teams indicated that for 37 ICD (1.3 percent), discovery was made by client and auditor jointly. Assuming that the clients might have missed these ICD without auditor involvement, we classify them with auditor-detected ICD.

⁹ Asking auditors to code associated misstatements was suggested by personnel of one of the participating firms, and so this question was asked of some but not all engagement teams. For remaining engagements, the authors identified associated misstatements from the text of the ICD description. If, for those engagements, the text did not mention a related misstatement when one existed, then the ICD would have been classified as not misstatement-associated. This should bias against finding differences between subsamples with/without associated misstatements.

¹⁰ We classify control flaws related to IT general controls (ITGC) based on AS No. 2 and guidance of the IT Governance Institute (2004) as system security, operations, development, and change process controls. General ledger issues relate to the following processes: financial statement close process, adjustments and accruals, allocations (e.g., depreciation and amortization), financial statement preparation, nonroutine entries, general ledger maintenance, SEC filings, and foreign currency.

AS No. 2: *DESIGN* (the control does not exist or cannot meet its objective); *DOCUMENTATION* (the control exists but is not documented); *INEFFECTIVE* (the control exists but does not function); *INSUFFICIENT_TESTING* (the control exists, but has not been adequately tested, perhaps due to insufficient time prior to year-end).¹¹

Table 2 provides names and definitions for company-level control variables, based on prior research showing that MW disclosure is associated with company size, financial condition, complexity, and earnings quality (e.g., Ashbaugh-Skaife et al. 2007; Doyle et al. 2007).¹² We measure company size using the natural log of total assets (*LNASSETS*), and financial condition using total liabilities divided by total assets (*LEVERAGE*) and cash flow from operations (*CFOTA*). We also

TABLE 2
Sample Composition and Descriptive Statistics on Company-Level Control Variables
Mean (Std. Dev.) or Percent = 1

	All Engagements (n = 76)	Effective Controls (n = 61)	Ineffective Controls (n = 15)
<i>LNASSETS</i> = natural log of assets.	19.52 (1.16)	19.54 (0.95)	19.43 (1.21)
<i>LEVERAGE</i> = total liabilities/total assets.	0.43 (0.22)	0.44 (0.22)	0.40 (0.17)
<i>CFOTA</i> = cash flow from operations/total assets.	-0.150 (0.19)	-0.012 (0.20)	-0.28 (0.24)
<i>ACCRUALS</i> = residuals from a regression of current-year cash flows on net income, within industry categories.	0 (0.09)	-0.007 (0.10)	0.027 (0.05)
<i>ARINVT</i> = (accounts receivable + inventory)/total assets.	0.29 (0.18)	0.27 (0.18)	0.36 (0.18)**
<i>SEGMENTS</i> = number of business segments.	2.51 (2.32)	2.38 (2.01)	3.07 (3.35)
<i>LOCATIONS</i> = number of significant locations.	4.87 (5.10)	4.84 (5.37)	5.00 (3.95)
<i>NEW_PUBLIC</i> = 1 if client has been a public company less than five years, and 0 otherwise.	30.3%	24.6%	53.3%**
<i>INDUSTRY_MFG</i> = 1 if client's industry is in the manufacturing sector (one-digit SIC = 2 or 3), and 0 otherwise.	56.6%	55.7%	60.0%
<i>FYE_2004</i> = 1 if fiscal year is 2004, and 0 otherwise.	50.0%	47.5%	60.0%

** Indicates ≤ 0.05 , significant effects in two-tailed tests.

This table presents descriptive statistics for company-level control variables, for the entire sample and for engagements with effective/ineffective ICFR. We test differences between companies with effective/ineffective ICFR using t-statistics for continuous variables and χ^2 tests (d.f. = 1) for dichotomous variables.

¹¹ To assure coding reliability, we used a practice sample of 73 ICD in two engagements, identified issues that reduced agreement, and refined our coding rules. We then independently coded 625 ICD from four engagements, with 97.2 percent agreement. The remaining ICD were coded by one or the other author.

¹² To protect client identities, financial information was disguised by engagement teams through multiplying by a factor between 0.95 and 1.05.

construct a measure of earnings quality by regressing net income on cash flow from operations within industry categories. The residuals from this regression constitute the variable *ACCRUALS*.¹³ We measure complexity as the ratio of accounts receivable and inventory to total assets (*ARINVT*), the number of reportable business segments (*SEGMENTS*), and the number of significant locations (*LOCATIONS*). We also include indicator variables for companies that have been public fewer than five years (*NEW_PUBLIC*), manufacturing companies (*INDUSTRY_MFG*),¹⁴ and fiscal year using *FYE_2004*.

Models

We address our first research question in two ways: (1) by examining the overall percentage of ICD detected by clients/auditors; and (2) by modeling factors associated with likelihood of client detection. The latter analysis uses Model 1, a logistic regression model whose dependent variable is *DETECTED_CLIENT*, which equals 1 if the ICD was detected by the client, and 0 if detected by the auditor or auditor/client jointly.¹⁵ Because we expect that clients are less likely to detect severe ICD, we expect negative coefficients on indicator variables *SD*, *MW*, and *ENTITY*. Because greater knowledge and independence from management of the parties performing the work should result in higher likelihood of client detection, we expect positive signs on *CONSULTANT_BIG6*, *CONSULTANT_OTHER*, *REPORT_INDEPENDENT*, and *INTERNAL_AUDIT*. Companies with better *IT_INTEGRATION* might perform better in ICD detection, but have better underlying controls, so we do not predict a sign on that variable. We expect a positive sign on *MONTHS_PRIOR_CLENT*, as a longer Section 404(a) process should help clients detect more ICD. We also expect a positive sign on *LNASSETS*, as larger companies should have more resources to devote to the process.

We investigate severity classification using two methods. First, because auditors are sometimes aware of the client's preliminary classification of ICD, we first test whether auditors override those classifications by judging ICD to be more severe. Second, we use two logistic regression models to test expectations regarding factors associated with severity classification of ICD, following the two-stage process implied in the definitions of *SD* and *MW* in AS No. 2 and explicitly featured in the profession's white paper (BDO Seidman LLP et al. 2004).¹⁶ Model 2 (whose dependent variable is *SD/MW*; 1 = *MW* or *SD*, 0 = Deficiency), investigates factors associated with ICD meeting the criterion for a "more than remote" likelihood of failing to detect or prevent a misstatement (i.e., at least *SD* classification, implying that the ICD must be reported at least to management and the audit committee). Model 3 is tested on the subsample of ICD having met the likelihood criterion, investigating factors associated with the auditor's judgment that misstatements resulting from the ICD could be material (which determines public disclosure of ineffective ICFR). The dependent variable of Model 3 is *MW* (1 = *MW*, 0 = *SD*).

In Models 2 and 3, we expect more severe ICD among clients with lower control quality. Control quality is measured using *IC_RELIANCE* and the number of *MW* per engagement (due to

¹³ Data limitations prevent us from estimating models to separate normal from abnormal accruals.

¹⁴ Our sample comprises public entities in non-regulated industries. Most sample companies are in manufacturing (SIC codes ranging from 2000 to 3999; 56.6 percent). Because for 19 of our 76 engagements, engagement teams did not provide a SIC code, we construct an indicator variable comparing manufacturing to all other industries.

¹⁵ Because there are multiple ICD for each company, all models are estimated with clustering on engagement ID (i.e., company/year), producing robust standard errors. Clustering is an important feature of our empirical tests, but it causes estimation difficulties when using some variables in smaller subsamples, due to insufficient numbers in one or both categories of the dependent variable. In that case, we omit as few variables as possible to enable estimation.

¹⁶ Use of two models allows coefficients to vary across the two stages. To investigate whether this approach is statistically indicated, we estimated an ordered probit with a three-level dependent variable, and applied the Brant test for proportional odds. This test is significant ($p < 0.01$), indicating that the parallel regression assumption is violated and the two-model approach is appropriate.

skewness, we use *NUM_MW_SQRT*, the square root of the number of MW detected). We also investigate the association of ICD severity classification with client process measures.¹⁷ We do not predict a sign on either *CONSULTANT* variable, as consultants may detect more severe control flaws, or enable earlier identification and remediation of severe flaws. We expect the sign on *REPORT_INDEPENDENT* to be positive, as it implies that Section 404 results are not filtered by management's incentives to classify ICD as less severe. We do not propose a sign on *INTERNAL_AUDIT*. Companies with an internal audit function might have less severe control flaws prior to Section 404 detection processes, and might also be better able to recognize severity than client personnel without internal audit experience. The sign on *IT_INTEGRATION* is also indeterminate. Companies with effective IT integration in the Section 404 process could perform better at detecting severe ICD. However, those companies likely also have better IT controls in general, meaning fewer severe problems should exist.

Models 2 and 3 include four auditor processes variables. We expect positive coefficients on *DETECTED_AUDITOR* (given auditors' greater knowledge and independence) and on *MONTHS_PRIOR_AUDITOR* (a longer auditor testing period provides more time to detect serious control flaws and consider their implications for financial reporting). We expect a positive sign on *MISSTATEMENT*, as a current misstatement provides objective evidence for assessing the likelihood and materiality of future misstatements. We also expect higher severity classifications for ICD detected with substantive tests (*DETECTED_SUBS*).

Regarding ICD type variables, we expect that entity-level ICD will be more severe, with one exception. Professional guidance during our study period (BDO Seidman LLP et al. 2004) indicates that regardless of the criteria of AS No. 2, auditors should classify ITGC consistent with underlying process controls. Thus, we do not propose a directional expectation on *ENTITY_ITGC*. Also, relative to the ICD in control documentation that remain in the intercept, we expect positive signs on *DESIGN*, *INEFFECTIVE*, and *INSUFFICIENT_TESTING*, as documentation problems should be more easily remediated than other types of control flaws.

We derive our predictions for company-level control variables from prior research on factors associated with public disclosure of MW (e.g., Doyle et al. 2007; Hoitash et al. 2009). We expect a negative sign on *LNASSETS*, as prior research shows smaller companies have more serious internal control problems. We expect that poor financial health will be associated with more severe ICD, generating a positive sign on *LEVERAGE* and a negative sign on cash flow from operations (*CFO*). We expect an association between ICD severity and *ACCRUALS*, based on prior research showing earnings quality to be lower among companies with internal control flaws (e.g., Ashbaugh-Skaife et al. 2009). We also expect that company complexity will be positively associated with ICD severity, including *ARINVTA*, *SEGMENTS*, and *LOCATIONS*. More severe ICD should be found among new public companies (*NEW_PUBLIC*). These models also control for industry membership using *INDUSTRY_MFG*, and for fiscal year using *FYE_2004*.

IV. RESULTS

Description of the Sample

The sample comprises 3,990 specific ICD detected in 76 engagements for 44 companies. An auditor's Section 404 opinion of ineffective ICFR was given for 15 sample engagements (19.7

¹⁷ Companies' choices of Section 404 processes (e.g., to use a Section 404 consultant) could be endogenous. Prior research does not address factors associated with these choices; thus, there is little guidance in the literature on what measures could be used as instruments. Available variables within our sample are few, and none fit the criteria for effective instruments: a strong theoretical link, statistical association with the selection variable, and no association with the outcome variable (severity classification). In such a situation, Larcker and Rusticus (2010) advise against formal modeling of selection, as greater bias can be introduced by use of poor instruments.

percent).¹⁸ Untabled analysis shows that an additional 51.4 percent of sample companies have at least one SD, implying that only 28.9 percent had no ICD of sufficient severity to be reported, at least to the audit committee. At the individual control level, 153 are MW (3.8 percent of detected ICD), a mean (median) of 10.2 (2.0) MW per company with ineffective controls. Most companies with ineffective controls have few MW, but some have a large number of ICD that auditors identified as MW or components of MW. However, this does not imply that each of these individual control flaws would be disclosed separately in Section 404 reports. Such disclosures vary in their level of specificity; for instance, individual deficiencies might be aggregated for discussion in the text of the opinion (e.g., several control flaws in revenues or several issues relating to the Control Environment). SD comprise 11.6 percent of sample ICD, and Deficiencies 58.4 percent, while 25.7 percent were remediated as of the balance sheet date. Details on ICD type by severity level are shown in Table 3.¹⁹

Table 2 presents descriptive statistics for company-level variables, and tests differences between subsamples with effective versus ineffective controls. The mean log of total assets is 19.52 (the mean in dollars is \$502.5 million), mean leverage is 0.43, and mean *CFOTA* is -0.150. The mean of *ACCRUALS* (the residual of regressing net income on cash flow from operations) is zero, by construction. The mean ratio of *ARINVTA* is 0.29; this ratio is larger for companies with ineffective ICFR (0.36 versus 0.27, $p = 0.048$). On average, sample companies have 2.51 segments and 4.87 locations. The overall percentage of companies that have been public fewer than five years is 30.3 percent; companies with ineffective controls are significantly more likely to be newly public (24.6 versus 53.3, $p = 0.030$). Over half of the companies (56.6 percent) are in the manufacturing industry.

Descriptive Statistics on ICD Detection

Table 1, Panel A describes company-level Section 404 process variables. Panel A shows that control reliance is higher in engagements with effective controls (54.1 versus 26.7 percent, $p = 0.057$). Consultants from large auditing firms were used for Section 404 activity by 13.2 percent of sample companies, while 63.2 percent used a smaller auditing firm or another provider.²⁰ Results of the client's Section 404 activity were reported independent of management in 34.2 percent of the sample. Thus, in about two-thirds of sample companies, Section 404 results were filtered through management before being made available to the auditor. Just under half of the companies (46.1 percent) have an internal audit unit. Companies with ineffective controls have a lower mean *IT_INTEGRATION* than those with effective controls (2.93 versus 3.39 on the scale of 1 = highly ineffective to 5 = highly effective, $p = 0.035$). Clients (auditors) began control tests an average of 5.82 (4.21) months before year-end.²¹ Panel A also shows that the mean number of MW among engagements with ineffective controls is 10.20, while the median is 2.01, indicating a

¹⁸ Bedard et al. (2007) note that the overall proportion of smaller accelerated filers with ineffective controls under Section 404 is 20 percent in 2004 and 14 percent in 2005. Our sample proportions of companies with ineffective ICFR in 2004 (2005) are 23.7 (15.8) percent, not significantly different from the Bedard et al. (2007) disclosure proportions ($p = 0.345$ and $p = 0.444$, respectively).

¹⁹ For efficiency, we do not present a matrix of correlations among independent variables. Estimating the severity models using regression to obtain variance inflation factors, we find that the highest is 4.42 (4.41) for *NUM_MW_SQRT* in Model 2 (3), lower than the level of 5 often used to indicate concern for multicollinearity.

²⁰ Untabled statistics show that the consultant managed the work in 35.1 percent of companies using a consultant, performed controls documentation in 68.4 percent, and controls testing in 86 percent.

²¹ There were numerous reports in the business press in the first year of Section 404 implementation of concern that some companies delayed their start of testing too long to do an effective job. In supplemental analysis, we investigate whether the client's Section 404 testing began earlier for 2005 engagements in our sample than for 2004 engagements. In 2004, client testing began 5.28 months before year-end, while in 2005 client testing began 6.34 months prior to year-end. This difference is significant ($t = 2.087$, $p = 0.020$).

TABLE 3
Percentage Composition of Detected ICD by Severity Category

Panel A: Entity-Level ICD

	All (n = 3,990)	Deficiencies (n = 2,328)	Significant Deficiencies (n = 461)	Material Weaknesses (n = 153)	Remediated (n = 1,027)
ENTITY_MONITORING = COSO category of Monitoring.	12.6%	12.0%	14.3%	11.2%	13.1%
ENTITY_CONTROLENVIR = COSO category of Control Environment.	7.2	8.4	6.3	9.2	4.8
ENTITY_INFOCOM = COSO category of Information/Communications.	1.7	1.7	0.9	2.0	2.4
ENTITY_RISKASSESS = COSO category of Risk Assessment.	0.3	0.4	0	0.7	0.1
ENTITY_GENERAL_LEDGER = controls over general ledger processes, including accruals, closing, consolidations, and SEC filings.	22.0	19.9	27.1	26.3	24.1
ENTITY_ITGC = information technology general controls.	21.4	21.9	26.3	4.6	20.9
Total entity-level controls	55.7%	55.6%	62.0%	46.1%	55.2%

Panel B: Type of Control Failure

	All (n = 3,990)	Deficiencies (n = 2,328)	Significant Deficiencies (n = 461)	Material Weaknesses (n = 153)	Remediated (n = 1,027)
INEFFECTIVE = ineffective control.	41.5%	39.3%	43.6%	54.6%	43.1%
DOCUMENTATION = failure to document a control.	28.6	30.7	21.5	17.8	29.0
DESIGN = control design flaw.	25.2	23.8	34.7	25.7	24.1
INSUFFICIENT_TESTING = client failure to sufficiently test a control.	8.7	8.8	9.5	16.5	6.9

Panel C: Cycle/Account

	All ICD (n = 3,990)	Deficiencies (n = 2,328)	Significant Deficiencies (n = 461)	Material Weaknesses (n = 153)	Remediated (n = 1,027)
ACC_PURCHASES = purchase to payment cycle.	23.8%	23.1%	23.4%	21.6%	25.8%
ACC_REVENUES = revenue to cash cycle.	19.7	17.8	20.8	48.4	19.0
ACC_COMPENSATION = compensation, payroll and benefits.	8.9	9.0	10.0	2.6	9.3

(continued on next page)

Panel C: Cycle/Account

	All ICD (n = 3,990)	Deficiencies (n = 2,328)	Significant Deficiencies (n = 461)	Material Weaknesses (n = 153)	Remediated (n = 1,027)
ACC_FIXEDINTANG = fixed assets and intangibles.	5.5	6.5	2.8	2.0	4.9
ACC_TAXES = taxes.	3.1	2.8	6.3	3.9	2.1
ACC_LIABILITIES = commitments and contingencies.	2.2	2.0	1.7	3.9	2.6
ACC_TREASURY = treasury and investments.	2.5	2.0	2.2	3.3	3.8
Total account-specific controls	64.2%	62.2%	66.6%	80.9%	65.3%

This table presents characteristics of individual ICD, overall and by severity classification. Percentages are computed on numbers of ICD for which we have data on the relevant variable. In some columns, percentages add to greater than 100 due to multiple categorizations of individual ICD.

highly skewed distribution. The mean number of SD for companies with ineffective (effective) controls is 18.27 (3.07), a significant difference ($p = 0.030$). The overall mean number of Deficiencies is 30.63.

Table 1, Panel B shows that auditors detected 72.1 percent of ICD. This statistic addresses our first research question, showing that auditors detect a large proportion of ICD even though clients are both documenting and testing controls. This percentage differs among severity classes of unremediated ICD ($p = 0.001$); it is highest for MW (84.3 percent) and lowest for deficiencies (74.7 percent). Panel B also shows that the percentage of misstatement-associated ICD increases with severity ($p < 0.001$), as does the percentage detected with substantive tests ($p < 0.001$). The percentage of ICD detected by control tests decreases with severity ($p = 0.042$).

Table 4, Panel A provides further evidence on detection method by documenting the types of ICD detected by auditors with substantive (control) tests. Column 1 shows that 81.5 percent of unremediated entity-level ICD are initially detected through control tests, while only 15.1 percent are detected through substantive tests. Within entity-level categories, control test detection rates are above 80 percent for all except general ledger controls, which include adjusting and closing entries. Account-specific ICD are also more likely to be detected through control tests, with tax accounts having the lowest of the listed percentages (64.8 percent overall). The only ICD type for which substantive and control test detection is roughly equivalent is misstatement-associated (45.0 versus 44.1 percent). Panel B shows that the percentage of misstatement-associated ICD detected with control tests is lower than for other ICD ($p < 0.001$). Columns 2 and 3 show a pattern of decreasing control test detection and increasing substantive test detection with increasing severity. Panel B shows that overall, the percentage of ICD detected with control tests decreases with severity ($p < 0.001$). Overall, this analysis shows that while more severe ICD are less likely to be detected through control tests, control tests still provide initial evidence for most MW. These results based on control-level data cast doubt on the effectiveness of relying on misstatements discovered by auditor substantive procedures to detect key flaws in controls over financial reporting, as suggested by KS from company-level data.

Client ICD Detection and Under-Classification

Table 5 presents results from estimating Model 1, examining factors associated with client ICD detection, for all unremediated ICD and for SD/MW only, along with descriptive statistics on model variables. Results for all unremediated ICD show that clients are less likely to detect ICD ultimately classified as MW or SD (both at $p = 0.011$) as well as entity-level ICD ($p = 0.024$). Client detection is more likely for larger clients ($p = 0.005$) and those using a large audit firm consultant ($p = 0.001$). However, independent reporting chain, internal audit function, IT integration, and longer client testing process are not associated with better client detection. Among SD/MW only, a large audit firm consultant again improves detection ($p = 0.005$) and entity-level ICD are less likely to be detected by clients ($p = 0.038$). In sum, Model 1 shows that clients are less likely to detect more severe and pervasive ICD, but use of large audit firm consultants and greater resources improves client detection.

According to AS No. 2, auditors should be aware of ICD detected by clients before beginning work in a given area of the audit. Because the auditor may also know the client's preliminary severity classification, we compare client and auditor ICD classification in these situations. Table 6, Panel A shows that 71.4 (64.8) percent of client-detected MW (SD) were initially under-classified by clients. These percentages are both significantly different from chance ($p = 0.005$ and $p = 0.023$, respectively), implying that auditors tend to override client classifications in a significant number of cases. Client under-classification could be due to lack of knowledge of the impact of specific controls, or an attempt to influence the auditor's classification to a less severe level.

TABLE 4
Supplemental Analysis of Detection Method for Auditor-Detected ICD

Panel A: Percent Detected with Control (Substantive) Tests, by ICD Characteristics

	1. All Unremediated	2. SD and MW	3. MW
Overall	82.1 (14.3)	72.7 (22.8)	66.7 (28.7)
Entity Level	81.5 (15.1)	74.1 (22.3)	63.3 (33.3)
<i>ENTITY_CONTROLENVIR</i>	83.7 (9.9)	88.6 (2.9)	100.0 (0.0)
<i>ENTITY_MONITORING</i>	83.3 (14.5)	77.3 (21.2)	60.0 (40.0)
<i>ENTITY_GENERAL_LEDGER</i>	69.7 (26.2)	55.8 (39.5)	56.8 (37.8)
<i>ENTITY_ITGC</i>	93.3 (5.0)	96.6 (2.6)	80.0 (20.0)
<i>ENTITY_RISKASSESS</i>	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)
Account-Specific	79.1 (17.3)	66.1 (30.1)	64.4 (32.7)
<i>ACC_PURCHASES</i>	81.6 (14.9)	71.0 (23.4)	72.0 (20.0)
<i>ACC_REVENUES</i>	79.1 (16.5)	72.0 (25.0)	66.2 (30.8)
<i>ACC_COMPENSATION</i>	84.0 (13.3)	64.9 (35.1)	66.7 (33.3)
<i>ACC_TAXES</i>	64.8 (29.7)	46.9 (40.6)	40.0 (60.0)
Misstatement	45.0 (44.1)	30.6 (57.3)	35.9 (48.7)

Panel B: Tests of Significance of Difference in Detection Method by ICD Characteristics

	Percent Detected with Control Tests	$\chi^2_{df=1}$
MW and SD Deficiencies	72.7% 84.9%	39.5***
Misstatement Associated All other ICD	45.0% 88.2%	341.1***

*** Indicates significance at ≤ 0.01 .

This table presents supplemental analysis of detection method, among those ICD detected with auditor involvement, for which that information is available.

Panel A shows percentages of control (substantive) test discovery for all unremediated ICD (Column 1), SD/MW (Column 2), and MW (Column 3). Percentages may not add to 100 due to other possible methods of detection.

Panel B provides tests of detection method by severity, and between misstatement and no-misstatement subsamples.

TABLE 5
Descriptive Statistics and Model Results on Client Detection

	Percent of Unremediated ICD Detected by Clients Alone	Results of Model 1	All Unremediated ICD	SD/MW
Overall	23.9%	CONSULTANT_BIG6 (+)	1.306***	1.735***
No Section 404 consultant	17.2	CONSULTANT_OTHER (+)	0.240	0.613
Smaller audit firm or other consultant	22.5	REPORT_INDEPENDENT (+)	0.171	0.358
Big 6 firm consultant	37.2	INTERNAL_AUDIT (+)	-0.588†	0.157
Section 404 results reported through management	23.4	IT_INTEGRATION	-0.330	-0.335
Section 404 reporting chain independent of management	24.6	MONTHS_PRIOR_CLIENT (+)	-0.038	0.078
No internal audit function	26.3	MW (-)	-0.827**	-0.195
Internal audit function	20.7	SD (-)	-0.472**	
Deficiencies	25.3	ENTITY (-)	-0.348**	-0.707**
SD	19.1	LNASSETS (+)	0.408**	0.319
MW	15.7	FYE_2004 (-)	-0.517*	-0.329
Entity level	20.2	Constant	-7.574**	-6.272**
Not entity level	28.5	n	2,930	612
		Pseudo-R ²	0.065	0.082

* ** *** Indicate significant effects at ≤ 0.10 , ≤ 0.05 , and ≤ 0.01 , respectively, probability levels are one-tailed for directional expectations indicated in the table. Values in the cells are model coefficients.

† Indicates significance in the opposite direction to that predicted, at $p \leq 0.10$ (two-tailed).

This table presents descriptive statistics and results of estimating Model 1, a logistic regression model (estimated with clustering on engagement ID) of factors associated with client detection of ICD without auditor assistance. Independent variables are defined in Tables 1 and 2.

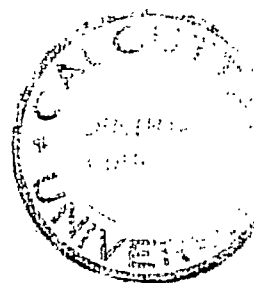


TABLE 6
Client Under-Classification of ICD Severity

Panel A: Overall Statistics on Client Under-Classification

	SD	MW
ICD detected by clients, for which the initial client classification is known	91	28
Number (percent) for which the initial client classification was "deficiency"	59 (64.8%)	16 (57.1%)
Number (percent) for which the initial client classification was "SD"		4 (14.3%)
Overall number (percent) under-classified	59 (64.8%)	20 (71.4%)
$\chi^2_{df=1}$	8.011***	5.143**

Panel B: Descriptive Statistics and Model Results on Factors Affecting Client Under-Classification of SD and MW

% of SD/MW Under-Classified by Clients	Logistic Regression Model of Under-Classification	
High/significant control reliance	CONSULTANT_BIG6 (-)	0.968
Other control reliance	CONSULTANT_OTHER (-)	1.214
No Section 404 consultant	REPORT_INDEPENDENT (-)	-18.603***
Smaller audit firm or other consultant	INTERNAL_AUDIT (-)	-0.107
Big 6 firm consultant	IT_INTEGRATION (-)	0.186
Reporting chain through management	MW (+)	17.875***
Reporting chain independent of management	ENTITY	-1.059**
No internal audit function	LNASSETS	-1.388
Internal audit function	FYE_2004	0.468
SD	Constant	27.990
MW	n	119
Entity level	Pseudo-R ²	0.513
Not entity level		

(continued on next page)

, *, *** Indicate significant effects at ≤ 0.05 and ≤ 0.01 , respectively, probability levels are one-tailed for directional expectations indicated in the table. Values in the cells are model coefficients.

This table presents supplemental analysis of client under-classification for those SD/MW for which a preliminary client classification is available. The model is a logistic regression estimated with clustering on engagement ID, whose dependent variable = 1 if the client's classification is lower than the auditor's final classification, and 0 if the same. We include descriptive statistics on *IC_RELIANCE*, but this variable is omitted from the model as its presence prevents estimation.

Table 6, Panel B presents supplemental statistics on client classification of ICD whose final classification is SD or MW, and results of estimating a logistic regression model whose dependent variable (*UNDER_CLASSIFICATION*) equals 1 if the client's preliminary classification is lower than the auditor's final classification, and 0 otherwise. We expect that clients might not recognize, or might seek to downplay, more severe (MW) or pervasive (entity-level) problems. Consistent with this expectation, the coefficient on *MW* is positive and significant ($p < 0.001$). However, entity-level ICD are unexpectedly less likely to be under-classified ($p = 0.035$, two-tailed).²² We also expect that clients with stronger Section 404(a) processes should be less likely to under-classify severity. Model results show that under-classification is less likely when the client has an independent reporting chain ($p < 0.001$), consistent with this expectation. Although Model 1 shows that large firm consultants improve ICD detection, Table 6 shows that independence in the reporting process plays a greater role in accurate client severity classification.²³

Results of Estimating Models 2 and 3: Auditors' Severity Classification Decisions

Table 7 presents results on factors affecting auditors' severity classification decisions, investigated using Model 2 (distinguishing SD/MW from Deficiencies) and Model 3 (distinguishing MW from SD). These models have a Pseudo- R^2 of 0.352 and 0.437, respectively.

Client Characteristics and Section 404 Processes

Model 2 results show a negative and significant coefficient on *IC_RELIANCE* ($p = 0.013$) and a positive and significant coefficient on *NUM_MW_SQRT* ($p < 0.001$). Both imply that, when the client's overall control quality is lower, an ICD is more likely to be classified as SD relative to a deficiency. Concerning Section 404 process strength, Model 2 shows a positive coefficient on use of a large accounting firm as consultant ($p < 0.001$), but not for other consultants.²⁴ SD/MW classification is also more likely when Section 404 reporting is independent of management ($p = 0.001$). The coefficient on *IT_INTEGRATION* in Model 2 is negative and significant ($p = 0.012$), implying that better underlying IT controls are associated with less severe ICD. The presence of an internal audit function is not significant. Model 3 results show factors that contribute to distinguishing MW from SD. Regarding the overall controls context, *NUM_MW_SQRT* is again significant ($p < 0.001$), but *IC_RELIANCE* is not. An independent reporting chain is also positive and significant in Model 3 ($p = 0.037$).

Auditor Section 404 Processes

Model 2 results show that as expected, ICD detected by auditors are more likely to be SD/MW ($p = 0.008$). The length of the auditor testing period is also positive and significant ($p = 0.003$), suggesting that auditors who take more time to consider consequences of ICD classify them as more severe. ICD detected with a current misstatement are more likely to be SD/MW ($p < 0.001$), as are those discovered through substantive tests ($p = 0.074$). In Model 3,

²² While Model 1 shows that clients are less likely to detect entity-level problems, this result implies they are more aware of their significance when detected.

²³ Although the descriptive statistics in Panel B include *IC_RELIANCE*, this variable is not included in the model, as its presence prevents estimation. This is likely due to the very low extent of under-classification among well-controlled companies. Table 5, Panel B shows that only 6.1 percent of ICD of clients of with higher control reliance are under-classified, relative to 89.5 percent of other clients.

²⁴ In Model 1 we find that large accounting firm consultants are associated with better ICD detection, and in Model 2 we find they are associated with higher severity classifications. These results could be explained by a tendency for poorly controlled companies to hire Big 6 consultants, but we find no association of use of large firm consultants with internal control reliance ($p = 0.555$). Model 2 also shows that an independent reporting chain is associated with higher severity classifications, which could be due to poorly controlled companies designing their processes in that manner. However, we also find no association of independent reporting with internal control reliance ($p = 0.869$). Thus, the more likely explanation is that use of those processes leads to higher quality Section 404(a) outcomes.

TABLE 7

Models 2 and 3: Control, Process, and Company Factors Associated with Severity Classifications

	Model 2: SD and MW versus Deficiencies	Model 3: MW versus SD
Client Characteristics and Detection Processes		
IC_RELIANCE (–)	–0.945**	0.086
TOTAL_MW_SQRT (+)	0.308***	0.763***
CONSULTANT_BIG6	1.892***	–1.473
CONSULTANT_OTHER	0.244	0.257
REPORT_INDEPENDENT (+)	1.464***	3.224**
INTERNAL_AUDIT	–0.217	2.266
IT_INTEGRATION	–0.527**	0.705
Auditor Detection Processes		
DETECTED_AUDITOR (+)	0.449***	0.240
MONTHS_PRIOR_AUDITOR (+)	0.324***	–0.680
MISSTATEMENT (+)	0.849***	0.636
DETECTED_SUBS (+)	0.559*	0.991*
ICD Types		
ENTITY_CONTROLENVIR (+)	0.434	2.603***
ENTITY_MONITORING (+)	0.228	–0.744 ^{††}
ENTITY_INFOCOM (+)	–1.612 ^{††}	
ENTITY_RISKASSESS (+)	–1.739 ^{†††}	
ENTITY_GENERAL_LEDGER (+)	0.425**	0.433
ENTITY_ITGC	0.445*	–1.113*
ACC_REVENUES (+)	0.933***	2.060***
ACC_TAXES (+)	0.851**	0.839
ACC_PURCHASES	0.392	0.811
ACC_COMPENSATION	0.191	
ACC_FIXEDINTANG	–0.755*	
ACC_LIABILITIES	–0.208	
ACC_TREASURY	0.496	
DESIGN (+)	1.183***	0.993***
INEFFECTIVE (+)	0.760***	0.928***
INSUFFICIENT_TESTING (+)	0.627***	0.817**
Control Variables: Client Characteristics		
LNASSETS (–)	–0.245	0.584
LEVERAGE (+)	–0.563	1.980
CFOTA (–)	3.715 ^{††}	–6.717**
ACCRUALS (+)	11.817***	–0.074
ARINVIA (+)	–4.915 ^{†††}	7.002**
SEGMENTS (+)	–0.249 ^{†††}	0.031
LOCATIONS (+)	–0.009	–0.232
NEW_PUBLIC (+)	0.848***	3.777**
FYE_2004 (–)	–0.361	–0.093
INDUSTRY_MFG	–0.979***	–1.488
Constant	3.655	–22.423
n	2,909	611
Pseudo-R ²	0.352	0.437

(continued on next page)

TABLE 7 (continued)

*, **, *** Indicate significant effects at ≤ 0.10 , ≤ 0.05 , and ≤ 0.01 , respectively, probability levels are one-tailed for directional expectations indicated in the table. Values in the cells are model coefficients.

††, ††† Indicate significance in the opposite direction to that predicted, at $p \leq 0.05$ and $p \leq 0.01$, respectively (two-tailed).

This table shows results of estimating Models 2 and 3, logistic regressions estimated using clustering on engagement ID. Model 2 explains factors differentiating deficiencies from SD (i.e., the ICD meets criteria for disclosure at least to the audit committee), and Model 3 explains factors differentiating SD from MW (i.e., the ICD meets criteria for public disclosure). Independent variables are defined in Tables 1 and 2. A blank cell in Model 3 results indicates that a variable is not included because its presence prevents estimation, due to low representation.

we again find that substantive test detection is associated with MW classification ($p = 0.062$). However, an associated misstatement does distinguish MW from SD, when controlling for the nature of the ICD and other variables in the models.

ICD Types

Our models also examine severity classifications of specific types of entity-level ICD. Model 2 results show differing associations across types. Coefficients on general ledger controls ($p = 0.036$) and ITGC ($p = 0.092$) are positive and significant. However, COSO categories of Information/Communications and Risk Assessment are negative and significant ($p = 0.034$ and $p < 0.001$, respectively, two-tailed), and Control Environment is not significant. In contrast, Model 3 shows that location in the Control Environment is a strong factor in distinguishing MW from SD ($p = 0.001$). Monitoring ICD are less likely to be MW than SD ($p = 0.054$), as are ITGC ($p = 0.085$).

We also examine the severity of specific account-level ICD types. Both revenue and tax accounts have positive coefficients in Model 2 ($p = 0.002$ and $p = 0.026$, respectively), consistent with regulators' attention to those areas prior to and during the study period. The revenue category is also significant in Model 3 ($p < 0.001$), likely due to the high volume of dollars flowing through that account. Model 2 results on the nature of detected ICD show positive and significant coefficients for design flaws ($p < 0.001$), ineffective controls ($p < 0.001$), and insufficient testing ($p = 0.003$). In Model 3, coefficients are again positive and significant for design flaws ($p = 0.003$), ineffective controls ($p < 0.001$), and insufficient testing ($p = 0.047$).

Control Variables

Models 2 and 3 also contain control variables for client characteristics. Model 2 results show that SD/MW are more likely for companies with lower earnings quality as measured by higher accruals ($p < 0.001$), and for new public companies ($p = 0.005$). SD/MW are unexpectedly less likely in more complex companies as measured by segments ($p < 0.001$) and concentration of assets in receivables and inventory ($p < 0.001$). Model 3 shows that MW classification is more likely in firms with lower cash flows ($p = 0.035$), greater complexity as measured by *ARINVT*A ($p = 0.047$), and in new public companies ($p = 0.030$).

Severity Classification with versus without an Associated Misstatement

Because prior research suggests differences in auditor judgment processes based on objectivity of evidence, we investigate the role of existing misstatements in severity classification by

estimating Models 2 and 3 in subsamples of ICD with/without an associated misstatement.²⁵ Table 8 compares coefficients on Section 404 process and ICD type variables in Models 2 and 3, across subsamples. We focus our discussion on variables with a significant cross-model comparison, and those that are significant in both models. In Model 2, all significant differences between subsamples are in the expected direction (i.e., a stronger association in the no-misstatement subsample). *NUM_MW_SQRT* is significant in both subsamples ($p = 0.059$ and $p < 0.001$, respectively), but the coefficient is statistically stronger in the no-misstatement subsample ($p = 0.029$). Thus, the client's ICFR context plays a larger role in severity classification when a misstatement has not occurred. Section 404 process variables are significant only in the no-misstatement subsample: *REPORT_INDEPENDENT* ($p = 0.002$), with the difference across models significant at $p = 0.010$, *IT_INTEGRATION* ($p = 0.003$), and *MONTHS_PRIOR_AUDITOR* ($p = 0.004$). Differences in coefficients for these three variables are significant across models with probabilities of at least $p < 0.05$. No ICD type variable is significantly different across subsamples, but both models show positive coefficients for revenues (misstatement: $p = 0.097$; no misstatement: $p < 0.001$) and design flaws (misstatement: $p = 0.022$; no misstatement: $p < 0.001$), indicating the pervasive importance of these ICD types.

Table 8 also shows results of comparing coefficients of Model 3 between misstatement and no-misstatement subsamples. *NUM_MW_SQRT* is again associated with MW classification in the no-misstatement subsample ($p = 0.004$) and its association is statistically greater than when there is a misstatement ($p = 0.010$). The same pattern holds for *REPORT_INDEPENDENT* ($p = 0.017$) and *DETECTED_SUBS* ($p = 0.002$). However, *IC_RELIANCE* is significant only when there has been a misstatement ($p < 0.001$), and the association is stronger in that subsample ($p = 0.010$). The latter result implies that when a misstatement amount is available, the auditor is less likely to judge the implications of the ICD to be material in clients for which the audit engagement was performed with the perception that overall controls were good. However, control reliance is not associated with MW classification when there is not a misstatement.²⁶ Among ICD type variables, revenue ICD are more likely to be MW in both subsamples ($p < 0.001$ and $p = 0.015$, respectively). General ledger ICD are significant only in the no-misstatement subsample ($p = 0.001$), and the test of association between subsamples is significant ($p = 0.039$).

Taken together, Table 8 results show that client and auditor process measures and characteristics are more likely to influence severity classification when the control has not yet failed to prevent/detect a misstatement. This pattern of findings is consistent with prior research on the impact of imprecise standards (Mayhew et al. 2001; Nelson et al. 2002; DeZoort et al. 2003), and implies that when a control failure is not yet documented, auditors use a wider variety of other information about the control flaw and its context to make and justify severity decisions.

V. CONCLUSIONS AND LIMITATIONS

This study contributes to the literature on ICFR by examining the relative effectiveness of companies and their auditors in Section 404 ICD detection, the processes used by those parties,

²⁵ Due to the difficulty of interpreting interactions in logit, we compare coefficients between subsamples using seemingly unrelated estimation in Stata. Our cross-model comparisons are limited in that the presence of some variables prevents estimation in certain subsamples (e.g., consultant use in Model 3).

²⁶ There are several possible reasons for this pattern, which our data cannot distinguish. One is that clients with high control reliance might have more effective compensating or monitoring controls, which would mitigate a MW classification. This could be due to auditors anchoring on the available misstatement amount (not on the potential amount of future misstatements) and judging that amount to be less severe for a client with good controls. A further reason to downplay the severity of an ICD with a detected misstatement for an engagement in which controls have been relied upon, is that if a MW is found, the whole audit strategy of reducing substantive testing due to good controls is called into question.

TABLE 8
Comparison of Model Coefficients in Subsamples With/Without an Associated Misstatement

	Model 2: (SD/MW versus Deficiencies)			Model 3: (MW versus SD)		
	Misstatement Subsample	No Misstatement Subsample	Cross-Model Comparison ($\chi^2_{df=1}$)	Misstatement Subsample	No Misstatement Subsample	Cross-Model Comparison ($\chi^2_{df=1}$)
IC_RELIANCE (-)	-0.727**	-0.808	0.00	-5.580***	2.264	9.80***
TOTAL_MW_SQRT (+)	0.134*	0.351***	3.61**	0.060	1.517***	5.43***
CONSULTANT_RIG6	0.879	1.913***	1.26	-5.178		
CONSULTANT_OTHER	-0.616	0.295	3.06*	-3.370		
REPORT_INDEPENDENT (+)	0.027	1.950***	5.48**	-1.409	6.003**	2.21*
INTERNAL_AUDIT	-0.687*	-0.113	0.64			
IT_INTEGRATION	0.287	-0.774***	7.86***	-0.285	2.091	0.35
DETECTED_AUDITOR (+)	-0.098	0.499***	0.81	-0.162	0.071	0.17
MONTHS_PRIOR_AUDITOR (+)	-0.002	0.371***	3.75**	0.472	-0.183	0.23
DETECTED_SUBS (+)	0.652**	0.391	0.19	0.049	3.762***	8.66***
ENTITY_CONTROLENVIR (+)	-0.183	0.542	0.42		3.661***	
ENTITY_MONITORING (+)	0.431	0.242*	0.13	0.183	-1.496†††	1.94
ENTITY_GL (+)	0.236	0.429*	0.12	-0.328	1.512***	3.10**
ENTITY_ITGC (+)		0.506**				
ACC_REVENUES (+)	0.615*	1.140***	0.68	2.974**	2.822***	0.02
ACC_TAXES (+)	1.461**	0.360	1.85			
DESIGN (+)	1.047**	1.102***	0.02	0.874	0.752**	0.00
INEFFECTIVE (+)		0.594***			0.709***	
INSUFFICIENT_TESTING (+)		0.621***			1.263***	

*, **, *** Indicate significant effects at ≤ 0.10 , ≤ 0.05 , and ≤ 0.01 , respectively, probability levels are one-tailed for directional expectations indicated in the table. Values in the cells are model coefficients.

This table shows partial results of estimating Models 2 and 3 in sub-samples of ICD with/without an associated misstatement.

(The complete models are estimated but for efficiency, only the test variables are tabled.) Independent variables are defined in Tables 1 and 2.

and factors influencing ICD severity classification. Our data provide the only systematic control-level evidence available on the baseline set of detected ICD. Of those, about one quarter were remediated prior to year-end, showing progress in improving controls not visible to the public. Of all ICD detected, fewer than 4 percent are publicly disclosed as MW, but many companies having a “clean” Section 404 report have at least one SD—a control flaw serious enough to report to the audit committee. These results demonstrate the importance of looking beyond publicly available information to capture the full picture ICFR problems detected under Section 404.

Prior research does not differentiate ICD detection by auditors and clients (Leone 2007). Our data show that auditors detect almost three-fourths of ICD, especially those that are more severe and pervasive. Importantly, this low level of client detection occurs when clients are aware that auditors will soon follow with their own tests. Our results also reveal company characteristics associated with better ICD detection, including use of large audit firm consultants, Section 404(a) reporting independent of management, and greater resources (company size). While we do not find the mere presence of an internal audit function improves detection, Lin et al. (2011) find that characteristics of internal auditors do influence the quality of companies’ Section 404 activities.

Our results also shed light on auditor detection processes, a controversial issue at the heart of the debate on the value of Section 404(b). We find that two-thirds of auditor-detected MW are signaled through control tests. Therefore, without auditor control testing, many key flaws in ICFR could have been missed (e.g., over 80 percent of ICD in Control Environment and revenues). KS present a contrasting perspective, based on Audit Analytics data indicating that Section 404 opinions of companies with ineffective controls often mention misstatements. KS therefore suggest that substantive procedures, enhanced by “tracing back” from misstatements to the underlying control failure, could be an effective surrogate for Section 404 testing in identifying most companies with ineffective controls.²⁷ There are several reasons underlying the difference in conclusions between their research and ours. Audit Analytics data are organized at the company level, not at the level of the individual MW. Thus, KS’s results reveal that misstatements are associated with at least one of a company’s MW, but not which or how many MW caused a misstatement.²⁸ Our data augment publicly available information by revealing the number of individual control flaws detected per company, and how each was detected. Thus, we show the degree to which a policy not relying on auditor testing of controls would miss important control flaws. Our data show that, for 27 percent of our companies with ineffective ICFR, no MW are associated with a misstatement, and in an additional 40 percent, one or more MW are not associated with a misstatement. This implies that many MW would likely be missed if relying solely on detected misstatements to identify control deficiencies. Our data additionally show that most SD would not be identified through detected misstatements. In sum, we conclude that relying on misstatements to identify ICD will result in missing many severe control flaws. Thus, the intended overall purpose of SOX Section 404 (i.e., to improve financial reporting) would not be completely fulfilled for those companies.

In addition to the difference in unit of analysis (company or ICD), our study differs from KS in the purpose of Section 404. They focus on disclosure of ineffective controls in the auditor’s opinion as that Section’s purpose. This view implies that if disclosure of ineffective controls (i.e.,

²⁷ Our data show a similar proportion of companies with ineffective controls for which at least one MW is associated with a misstatement. KS cite a range of 57 to 73 percent based on the aggregated data; in our sample this proportion is 73 percent.

²⁸ Additionally, reliance on “tracing back” from misstatements to identify related control flaws assumes that the yield of misstatements detected from substantive tests would be the same without the information provided by the preceding extensive control tests required under Section 404. This seems unlikely, given that control tests should direct the auditor’s attention to areas of greater risk.

at least one MW) was obtained for a company through an alternative approach, then a purpose of Section 404 would be fulfilled for that company. However, the purpose of Section 404 requirements is to identify and correct control deficiencies and improve the reliability of financial reporting (e.g., COSO 2006; PCAOB 2004). Thus, an ineffective controls opinion is a consequence of the process of controls assessment, but not its main purpose. In sum, publicly available data are limited in supporting a conclusion on the efficiency and effectiveness of the controls assessment and testing process in improving financial reporting reliability. For the reasons cited above, our data provide a more complete picture of the results of Section 404 activity.

We also study how auditors classify ICD in the context of imprecise standards. For auditors, this situation is analogous to imprecise financial reporting standards, which have been shown to decrease auditors' ability to prevail in booking proposed audit adjustments (Nelson et al. 2002). Our results show that ICD associated with an existing misstatement have higher severity classifications. It is likely that a current control failure makes future failures easier to visualize, and provides more objective evidence to justify severity to clients (e.g., Kennedy et al. 1997). Nevertheless, our data show that a misstatement is neither necessary nor sufficient for high severity classification, prompting our investigation of how auditors make severity judgments when the control has *not yet failed*. For these ICD, results generally imply a stronger association of task expertise and process independence, relative to those ICD that have failed in the current period. Use of a large accounting firm consultant, lower quality of internal control, better client IT integration, and auditor substantive test detection are more strongly associated with severity when the objective evidence provided by a current misstatement is unavailable.

Our investigation of the nature of ICD also contributes to the literature by showing which types of entity-level and account-level controls that auditors consider could produce future misstatements most threatening to financial reporting quality. Within entity types, it is interesting to compare our findings on COSO Control Environment and Risk Assessment categories. Relatively few Control Environment ICD are detected, but those are more likely to be classified as MW. In contrast, COSO Risk Assessment ICD are infrequently found, and are *less* likely to be classified as severe. While this might imply that companies' risk assessment controls were of high quality during the study period, evidence from the recent financial crisis suggests that many companies did not effectively monitor their risks. Further research should examine whether companies and auditors have increased their focus on the quality of risk assessment controls in the current environment.

Among specific accounts/cycles, revenue ICD are more likely to be severe, consistent with the importance of revenues in equity valuation. We also find that tax ICD are more likely to be classified as at least SD. This is consistent with auditors' concerns that many clients during the study period were unable to perform the tax accrual without assistance. Regarding the nature of control failures, we find that ineffective controls (well designed but not functioning) are most frequent. Those ineffective controls, as well as design flaws (comprising missing controls as well as controls that do not meet their intended purpose), and controls that have been insufficiently tested by clients, are all considered more severe relative to documentation problems. Further research should investigate whether the incidence of these more severe types of control flaws has declined following several years of ICD remediation and retesting under Section 404.

Our study complements a behavioral experiment by Earley et al. (2008), which finds that clients' classifications influence auditors' judgments. Our archival data show that clients' classifications are overwhelmingly lower than those of auditors, implying that auditors often override

clients' views.²⁹ Client under-classification could be caused by differing incentives, uncertainty regarding appropriate procedures (due to lack of precise SEC guidance for company managements during our study period), and/or lack of knowledge of internal controls. The first explanation seems consistent with our finding that clients with a more independent Section 404(a) process are less likely to under-classify, but future research should investigate these and other possible causes.

The generalizability of this study's findings may be limited by several features of its design. First, our conclusions are limited by restricting the sample to smaller accelerated filers that more closely apply to non-accelerated filers and foreign companies not currently subject to Section 404 or similar regulation. Second, engagements in our sample were performed under AS No. 2. Additional guidance on company and auditor processes for Section 404 were subsequently issued (SEC 2005; 2006), and AS No. 5 (PCAOB 2007) superseded AS No. 2. Both the revised auditing standard and related SEC guidance were directed toward helping companies and auditors achieve greater efficiencies but not to "lower the bar" for reporting issues. While the current study's results should have continuing relevance, research in the AS No. 5 environment would provide insight on effects of application of that standard. Third, it is difficult to directly compare the proportions of MW identified in our sample with public disclosures, as individual control flaws may be aggregated in those reports. Fourth, sample engagements were chosen by participating firms. While our sample represents publicly available data at the company level (as indicated by statistical comparisons with public data reported above), we cannot rule out the possibility that nonrandom selection might have influenced our results.

In sum, this study provides a unique view into processes and outcomes of the first two years of Section 404 compliance. In terms of process, our results support the value of auditor involvement at two stages of the ICFR assessment process (detection and classification), and contribute to the understanding of factors associated with client and auditor performance in both stages. Regarding outcomes, we show that even companies with a clean Section 404 auditor's opinion can have numerous and/or significant ICD. Having been identified, those ICD may now be corrected, thus improving financial reporting quality. Overall, results of this study imply that the recently enacted exemption of non-accelerated filers from SOX Section 404(b), which requires auditor testing and an audit report on internal control, is likely to undermine the objectives of SOX for these entities, resulting in relatively less reliable financial reporting than had 404(b) provisions been implemented. Further, risk assessments in subsequent years may be less effective without the detailed information on client controls that comes to light when the independent auditor detects and classifies control flaws.

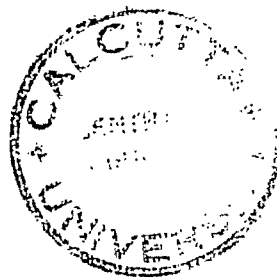
REFERENCES

- Allen, R., D. Hermanson, T. Kozloski, and R. Ramsay. 2006. Auditor risk assessment: Insights from the academic literature. *Accounting Horizons* 20 (2): 157–177.
- Amer, T., K. Hackenbrack, and M. Nelson. 1994. Between-auditor differences in the interpretation of probability phrases. *Auditing: A Journal of Practice & Theory* 13 (1): 126–136.
- American Institute of Certified Public Accountants (AICPA). 1988. *The Auditor's Consideration of Internal Control Structure in the Financial Statement Audit*. SAS No. 55. New York, NY: AICPA.
- . 1997. *Consideration of Internal Control in a Financial Statement Audit: An Amendment to Statement on Auditing Standards No. 55*. SAS No. 78. New York, NY: AICPA.

²⁹ Prior research finds lower rates of MW disclosure under Section 302, which does not require companies to test controls (e.g., Hoitash et al. 2009; Hermanson and Ye 2009). The current study shows that clients under-classify severity even when they test controls.

- Asare, S., and A. Wright. 2008. Equity analysts' reactions to type of control deficiency and likelihood threshold in adverse control reports. Working paper, University of Florida.
- Ashbaugh-Skaife, H., D. Collins, and W. Kinney. 2007. Discovery and reporting of internal control deficiencies prior to SOX-mandated audits. *Journal of Accounting and Economics* 44 (1-2): 166-192.
- , D. Collins, W. Kinney, and R. LaFond. 2008. The effect of SOX internal control deficiencies and their remediation on accrual quality. *The Accounting Review* 83 (1): 217-250.
- BDO Seidman LLP, Crowe Chizek & Company LLC, Deloitte & Touche LLP, Ernst & Young LLP, Grant Thornton LLP, Harbinger PLC, KPMG LLP, McGladrey & Pullen LLP, PricewaterhouseCoopers LLP, and W. Messier. 2004. *A Framework for Evaluating Deficiencies and Exceptions*. Version 3, December 20. New York, NY: BDO Seidman LLP et al.
- Bedard, J., R. Hoitash, and U. Hoitash. 2007. Audit pricing and internal control disclosures among non-accelerated filers. *Research in Accounting Regulation* 20: 103-126.
- Cohen, J., and D. Hanno. 2000. Auditors' consideration of corporate governance and management control philosophy in preplanning and planning judgments. *Auditing: A Journal of Practice & Theory* 19 (2): 133-146.
- , G. Krishnamoorthy, and A. Wright. 2002. Corporate governance and the audit process. *Contemporary Accounting Research* 19 (4): 573-594.
- Committee of Sponsoring Organizations (COSO). 2006. *Internal Control over Financial Reporting—Guidance for Smaller Public Companies*. New York, NY: AICPA.
- Curtis, M., D. Deis, G. Jenkins, and J. Bedard. 2009. Information systems and the financial statement auditor: A research synthesis. *Journal of Information Systems* 23 (1): 79-96.
- DeZoort, T., D. Hermanson, and R. Houston. 2003. Audit committee support for auditors: The effects of materiality justification and accounting precision. *Journal of Accounting and Public Policy* 22 (2): 175-199.
- Doyle, J., W. Ge, and S. McVay. 2007. Determinants of weaknesses in internal control over financial reporting. *Journal of Accounting and Economics* 44 (1-2): 193-223.
- Earley, C., V. Hoffman, and J. Joe. 2008. Reducing management's influence on auditors' judgments: An experimental investigation of SOX 404 assessments. *The Accounting Review* 83 (6): 1461-1485.
- Elder, R. A., and R. D. Allen. 2003. A longitudinal field investigation of auditor risk assessments and sample size decisions. *The Accounting Review* 78 (4): 983-1002.
- Felix, W., A. Gramling, and M. Maletta. 2001. The contribution of internal audit as a determinant of external audit fees and factors influencing this contribution. *Journal of Accounting Research* 39 (3): 513-534.
- Financial Accounting Standards Board (FASB). 1975. *Accounting for Contingencies*. Statement of Financial Accounting Standards No. 5. Stamford, CT: FASB.
- Hammersley, J., L. Myers, and C. Shakespeare. 2008. Market reactions to internal control weakness disclosures and to the characteristics of those weaknesses under section 302 of the Sarbanes-Oxley Act of 2002. *Review of Accounting Studies* 13 (1): 141-165.
- Hermanson, D., and Z. Ye. 2009. Why do some accelerated filers with SOX Section 404 material weaknesses provide early warning under Section 302? *Auditing: A Journal of Practice & Theory* 28 (2): 247-271.
- Hoitash, R., U. Hoitash, and J. Bedard. 2008. Audit fees and internal control problems disclosed under the Sarbanes-Oxley Act. *Auditing: A Journal of Practice & Theory* 27 (1): 105-126.
- Hoitash, U., R. Hoitash, and J. Bedard. 2009. Corporate governance and internal control quality: A comparison of regulatory regimes. *The Accounting Review* 84 (3): 839-867.
- IT Governance Institute. 2004. *IT Control Objectives for Sarbanes-Oxley: The Importance of IT in the Design, Implementation and Sustainability of Internal Control over Disclosure and Financial Reporting*. Chicago, IL: IT Governance Institute.
- Kennedy, J., D. Kleinmuntz, and M. Peecher. 1997. Determinants of the justifiability of performance in ill-structured audit tasks. *Journal of Accounting Research* 35 (Supplement): 105-123.
- Kinney, W.R., and M. Shepardson. 2010. Do control effectiveness disclosures require internal control audits: A natural experiment with small U.S. public companies. Forthcoming. *Journal of Accounting Research*.
- Krishnan, J., D. Rama, and Y. Zhang. 2008. Costs to comply with SOX Section 404. *Auditing: A Journal of Practice & Theory* 27 (1): 169-186.

- Larcker, D. F., and T. O. Rusticus. 2010. On the use of instrumental variables in accounting research. *Journal of Accounting and Economics* 49 (3): 186–205.
- Leone, A. 2007. Factors related to internal control disclosure: A discussion of Ashbaugh, Collins and Kinney (2007) and Doyle, Ge, and McVay (2007). *Journal of Accounting and Economics* 44 (1–2): 224–237.
- Li, C., L. Sun, and M. Ettredge. 2010. Financial executive qualifications, financial executive turnover, and adverse SOX Section 404 opinions. *Journal of Accounting and Economics* 50 (1): 93–110.
- Lin, S., M. Pizzini, M. Vargus, and I. Bardhan. 2011. The role of internal audit function in the disclosure of material weaknesses. *The Accounting Review* 86 (1).
- Mayhew, B., W. J. Schatzberg, and G. Sevcik. 2001. The effect of accounting uncertainty and auditor reputation on auditor independence. *Auditing: A Journal of Practice & Theory* 20 (2): 49–70.
- Messier, W., N. Martinov-Bennie, and A. Eilifsen. 2005. A review and integration of empirical research on materiality: Two decades later. *Auditing: A Journal of Practice & Theory* 24 (2): 153–187.
- Nelson, M., J. Elliott, and R. Tarpley. 2002. Evidence from auditors about managements' and auditors' earnings management decisions. *The Accounting Review* 77 (Supplement): 175–202.
- Ng, T., and H.-T. Tan. 2003. Effects of authoritative guidance availability and audit committee effort on auditor's judgments in an auditor-client negotiation context. *The Accounting Review* 78 (3): 801–818.
- Public Company Accounting Oversight Board (PCAOB). 2004. *An Audit of Internal Control over Financial Reporting Performed in Conjunction with an Audit of Financial Statements*. Auditing Standard No. 2. Washington, D.C.: PCAOB.
- . 2007. *An Audit of Internal Control over Financial Reporting that is Integrated with an Audit of Financial Statements*. Auditing Standard No. 5. Washington, D.C.: PCAOB.
- Securities and Exchange Commission (SEC). 1999a. *Materiality*. SEC Staff Accounting Bulletin No. 99. Washington, D.C.: SEC.
- . 1999b. *Revenue Recognition in Financial Statements*. SEC Staff Accounting Bulletin No. 101. Washington, D.C.: SEC.
- . 2005. Staff statement on management's report on internal control over financial reporting. (May 16). Washington, D.C.: SEC.
- . 2006. SEC announces next steps for Sarbanes-Oxley implementation. (May 17). Washington, D.C.: SEC.
- U.S. Congress. 2002. *The Public Company Accounting Reform and Investor Protection Act of 2002 (The Sarbanes-Oxley Act)*. Public Law No. 107-204, 116 Statute 745 (July 30). Washington, D.C.: Government Printing Office.
- . 2010. *Dodd-Frank Wall Street Reform and Consumer Protection Act*. Public Law No. 111-203 (July 21). Washington, D.C.: Government Printing Office.
- Wolfe, C., E. Mauldin, and M. Diaz. 2009. Concede or deny: Do management persuasion tactics affect auditor evaluation of internal control deviations? *The Accounting Review* 84 (6): 2013–2037.



Capital Structure, Cost of Capital, and Voluntary Disclosures

Jeremy Bertomeu
Northwestern University

Anne Beyer
Stanford University

Ronald A. Dye
Northwestern University

ABSTRACT: This paper develops a model of financing that jointly determines a firm's capital structure, its voluntary disclosure policy, and its cost of capital. Investors who receive securities in return for supplying capital sometimes incur losses when they trade their securities with an informed trader. The firm's disclosure policy and the structure of its securities determine the information advantage of the informed trader and, hence, the size of investors' trading losses and the firm's cost of capital.

We establish a hierarchy of optimal securities and disclosure policies that varies with the volatility of the firm's cash flows. Debt securities are often optimal, with the form of debt—risk-free, investment grade, or “junk”—varying with the firm's cash flow volatility. Though the model predicts a negative association between firms' cost of capital and the extent of information firms disclose, more expansive voluntary disclosure does not cause firms' cost of capital to decline. Mandatory disclosures alter firms' voluntary disclosures, their capital structure choices, and their cost of capital.

Keywords: *cost of capital; capital structure; voluntary disclosures.*

I. INTRODUCTION

We develop a model that jointly explains a firm's voluntary disclosure policy, its capital structure, and its cost of capital. While not previously recognized in the literature, it is intuitive that there should be a relationship between a firm's capital structure and its disclosure policy, because—barring agency problems—managers will choose their firm's disclo-

We thank seminar participants at the 2009 AAA Annual Meeting, Carnegie-Mellon University, Stanford University, University of Bern, University of California, Berkeley, University of California, Los Angeles, University of Pennsylvania, and Washington University in St. Louis, as well as Stan Baiman, Edwige Cheynel, Enrique de la Rosa, Pingyang Gao, Miles Gietzman, Jon Glover, Ilan Guttman, Jack Hughes, Pierre Liang, Ivan Marinovic, Paul Newman (editor), Tjomme Rusticus, Hyun Shin, Sri Sridhar, Phil Stocken, Amir Ziv, and two anonymous referees for comments on previous drafts of the manuscript. Professor Dye also thanks the Accounting Research Center at Northwestern University for financial support.

Editor's note: Accepted by Paul Newman.

Submitted: January 2009
Accepted: October 2010
Published Online: April 2011

sure policy so as to maximize the market's perceptions of the expected value of the firm-owners' residual claims. Hence, the form of those residual claims—equivalently, the form of the securities that the firm sells to investors—affects what managers maximize. In turn, the firm's disclosure policy and its capital structure choices affect the firm's cost of capital.

In the model we study, a firm receives capital by issuing securities that promise future cash payouts to investors. The central feature of the model is that the amount investors are willing to pay for the securities may be less than the expected present value of the securities' cash payouts because the investors anticipate at the time they buy the securities that they may bear trading losses if they have to liquidate the securities "early"—i.e., before maturity. These potential trading losses arise because the market in which the securities are traded is presumed to be populated by some superiorly informed traders. Compensating investors for these expected trading losses *ex ante* constitutes a cost the firm bears in raising capital: it is *the* source of the "cost of capital" in our model.

The firm's manager can reduce investors' expected trading losses by reducing the information asymmetry between the superiorly informed traders and the market maker who sets the price for the securities. This can be accomplished in two related ways. First, since the informed traders' private information about the firm is likely to overlap with the manager's private information, the manager can disclose his private information.¹ Second, the manager can have the firm issue securities whose value is not "informationally sensitive" (e.g., Sunder 2006), meaning that their value does not vary with the informed traders' private information. For example, were the securities risk-free debt, then the "informed" traders—regardless of their information—would have no information advantage over the market maker in assessing the securities' value. Thus, a firm's capital structure and its disclosure policy jointly determine how much information asymmetry remains among market participants, and this remaining information asymmetry determines investors' expected trading losses and, hence, the firm's cost of capital.

We develop a formula, applicable to all securities, that establishes how a firm's cost of capital depends on the security's upside potential (the difference between a security's payoff if the firm's realized cash flows assume "high" or "medium" values) and its downside risk (the difference between a security's payoff if the firm's realized cash flows assume "medium" or "low" values). The formula shows that the contribution of a security's upside potential and downside risk to the firm's cost of capital varies with the firm's disclosure policy. Specifically, the contribution of a security's upside potential to the firm's cost of capital decreases as the firm discloses more information, and, perhaps surprisingly, the contribution of a security's downside risk to the firm's cost of capital *increases* as the firm discloses more information.

Our model establishes that debt is always an optimal security and that there is a hierarchy among debt securities and optimal disclosure policies that varies with the volatility of the firm's cash flows. A firm with very low volatility in its cash flows prefers to raise capital by issuing risk-free debt and adopting an "expansive" disclosure policy.² As its cash flow volatility increases, the firm prefers to use investment-grade debt (that defaults with low probability) combined with the continued use of an expansive disclosure policy. Then, as its cash flow volatility increases still further, the firm will continue to use investment-grade debt, but it will curtail its disclosures and adopt a "limited" disclosure policy. As its cash flows become even more volatile, the firm will switch to using "junk" debt (that defaults if anything other than the highest cash flows occur),

¹ We conform to the AAA's style guidelines by mixing gender references in the following: specifically, when describing or referring to actors in the model, we arbitrarily assigned the gender of the manager and market maker to be male, and the gender of the investor and informed trader to be female.

² We defer the precise definitions of an "expansive" and a "limited" disclosure policy to the text below.

accompanied by limited disclosure. Finally, as its cash flow volatility becomes greater still, we demonstrate that it is impossible to finance the firm with any form of security accompanied by any form of disclosure.

While reminiscent of Myers and Majluf's (1984) famous financial hierarchy (inside financing is least expensive; outside financing with debt is more expensive; outside financing with equity is most expensive), our hierarchy is distinct from Myers and Majluf's (1984) hierarchy in two fundamental respects. First, our hierarchy combines a firm's capital structure choice with its voluntary disclosure policy, whereas Myers and Majluf (1984) make no reference to firms' disclosures. Second, our capital structure hierarchy is indexed to the volatility of a firm's cash flows, whereas Myers and Majluf's (1984) hierarchy is indexed to the relative costliness of alternative financing methods, holding cash flow volatility fixed.

We also explore how mandatory disclosure requirements interact with firms' voluntary disclosure decisions to affect firms' cost of capital and securities design choices. Some mandatory disclosure requirements are shown to inhibit firms from making voluntary supplementary disclosures, and other mandatory disclosure requirements are shown to encourage firms to make voluntary supplementary disclosures. In addition, we show that imposing mandatory disclosure requirements sometimes radically alters what securities firms wish to offer their investors. In particular, we show that it is sometimes optimal for the firm to raise capital by issuing equity rather than debt.

Related Literature

While links between a firm's disclosure policy and its cost of capital have been established in prior accounting research (e.g., Botosan 1997; Botosan and Plumlee 2002), and links between a firm's capital structure and its cost of capital separately have been identified through research in finance (e.g., in Myers and Majluf 1984), we believe this is the first study in accounting or finance that endogenously connects a firm's disclosure policy to its capital structure, and *a fortiori*, the first that connects all three of these components of a firm's financial structure.

Other models in accounting and finance have emphasized how information asymmetry affects firms' decisions to raise capital. For example, Korajczyk et al. (1991, 1992) noted that firms tend to raise new capital in periods where the information asymmetry between insiders of the firm and outside investors is small so as to eliminate the lemons' problem with issuing equity (noted originally by Myers and Majluf [1984]). Unlike Korajczyk et al. (1991, 1992), we are (1) concerned about not when firms raise capital, but rather in what form firms raise capital, and (2) not concerned about disclosures before the firm raises capital, but rather about disclosures after the firm has raised capital.³

Several recent studies discuss links between the quality of firms' financial statements and their cost of capital. For example, Jorgensen and Kirschenheiter (2003) show that firms have incentives to disclose information selectively that indicates lower cash flow risk, thus leading to low-risk premia conditional on voluntary disclosure. As another example, Christensen et al. (2002) develop a theory that explains when and how managers' compensation should be adjusted for the cost of capital.

Our model is also linked to the large literature in finance on securities design. For example, see Harris and Raviv (1991) for an early survey of this literature, and Nachman and Noe (1994), DeMarzo and Duffie (1999), Demange and Laroque (1995), Rahi (1995), Fulghieri and Lukin (2001), Sunder (2006), Bias et al. (2007), Casamatta (2003), and Repullo and Suarez (2004) for

³ But, the present model and Korajczyk et al. (1991, 1992) share the important property that the disclosure of information has real effects, which is also true of a substantial part of the disclosure literature (e.g., Kanodia et al. 2000; Stocken 2000; Pae 2002).

more recent work. As far as we are aware, none of the prior research on securities design has emphasized the interdependencies among securities design, cost of capital, and voluntary disclosures, which is our focus.

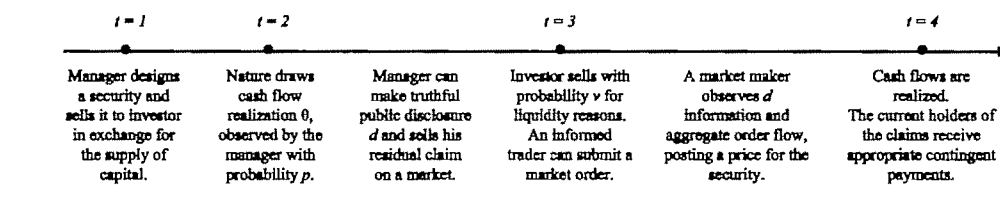
II. MODEL SETUP

The timeline for the model consists of four steps, presented in Figure 1. At step $t = 1$, an individual (called the “manager”) owns a production technology (called the “firm”) that converts capital invested in the first step stochastically into cash flows $\bar{\theta}$ in the fourth step. The manager, who has no capital, offers (one unit of) a security $\tilde{\psi}$ to an investor in return for the investor supplying capital to the firm.

At step $t = 2$, with probability $p \in (0,1)$ the manager privately learns the realization of the firm’s cash flows, and with probability $1 - p$, the manager learns nothing about those cash flows. If the manager receives no information, then he necessarily makes no disclosure. But, if he does receive information, then he may either disclose that information or disclose nothing, based on whether the disclosure or nondisclosure increases the capital market’s perceptions of the expected value of the manager’s residual claim $\bar{\theta} - \psi(\bar{\theta})$.⁴ We assume all disclosures are truthful, but the manager cannot credibly disclose that he has received no information. We let $d(\theta) = \emptyset$ indicate that the manager makes no disclosure and $d(\theta) = \theta$ indicate that the manager discloses that the cash flows are θ .

At step $t = 3$, a market for the security $\tilde{\psi}$ opens.⁵ With probability v , the investor experiences a liquidity shock that forces her to sell her entire interest in security $\tilde{\psi}$ on this market. We let $q = 1$ ($q = 0$) indicate that the investor is (is not) subject to a shock, so $\Pr(\tilde{q} = 1) = v$.⁶ If not subject to a shock, though the investor may still sell her security, it will be clear in what follows that she will elect not to. The market clearing price of security $\tilde{\psi}$ is then established via a Kyle (1985)-type market maker, who sets the security’s market price based on knowledge of the aggre-

FIGURE 1
Timeline



⁴ We take the assumption that the manager seeks to maximize the expected value of his residual claim as exogenous. In an earlier working paper, we endogenize the objective function. The working paper version also extends the technology to two required inputs, capital and “effort” supplied by the manager.

⁵ In a working paper version, we had two securities markets open, one for the security $\tilde{\psi}$ and one for the manager’s residual security $\bar{\theta} - \tilde{\psi}$. The results obtained by adding this second securities market are the same as in the present model.

⁶ Our results do not depend on the assumption of an asymmetry in the frequency of the investor’s and manager’s liquidity shocks. What is critical is not how frequently the manager faces a liquidity shock *vis-à-vis* the investor, but rather that, whether subject to a liquidity shock, the manager decides whether to disclose his private information (when he receives it) based on whether the disclosure increases his firm’s expected selling price.

gate order flow in the security. This aggregate order flow potentially includes orders from an informed trader. While the informed trader may go “long” in security $\tilde{\psi}$, we presume she is prohibited from “shorting” the security.⁷

Formally, we represent the informed trader as one who privately knows, when the manager makes no disclosure, whether the reason for the manager’s nondisclosure was due to the manager not having received information or having deliberately withheld information he did receive.⁸ In the latter case—where the investor knows that the manager withheld information—we presume that the trader does not know the value of the information the manager received. This constitutes the source of the informed trader’s informational advantage over the market maker; when the manager makes no disclosure, the market maker, unlike the informed trader, does not know the reason for the manager’s nondisclosure. This formal representation is intended to depict one of many possible ways an informed trader knows less than the manager. Results similar to those we obtain would hold for other representations where the informed trader’s information is a coarsification of the manager’s information.

At step $t = 4$, the firm’s cash flows occur, and the holders of the security at this time are paid off according to the terms specified by the security.⁹

Having completed the description of the timeline for the model, we conclude this introduction by offering more specifics about both the firm’s stochastic production technology and the security the manager offers investors. The production technology requires a fixed amount of capital, which we scale to be \$1, to be supplied by an investor. Less than \$1 of capital generates no cash flows in step 4 at all, and more than \$1 does not increase step 4’s cash flows. For notational convenience, we normalize the investor’s required rate of return on capital to 0. Given these normalizations, the manager can induce the investor to supply sufficient capital to run the technology by ensuring that the security $\tilde{\psi}$ has an expected payoff to the investor of at least \$1.

With the \$1 capital input, the firm’s cash flows (in step 4) are given by the realization of the random variable $\tilde{\theta}$, which assumes one of three possible values: $\theta_l \equiv \mu - e$, or $\theta_m \equiv \mu$ or $\theta_h \equiv \mu + e$, for some $e > 0$. Each of θ_l and θ_h occurs with probability $\frac{\eta}{2} \in (0, \frac{1}{2})$, and θ_m occurs with probability $1 - \eta$. We further assume that $\mu - e > 0$ and $\mu > 1$. In short, the firm’s realized cash flows are symmetrically distributed around their mean μ ; the cash flows are always positive, and investing in the production technology is efficient.

The security $\tilde{\psi}$ yields a payoff that is potentially contingent on the firm’s realized cash flows. Given that the firm’s cash flows can take on three values, a security $\tilde{\psi}$ can be described completely via the triple (ψ_l, ψ_m, ψ_h) , with $\psi_i = \psi(\theta_i)$. The securities we study include debt, equity, hybrids that mix debt and equity, and, more generally, any security $\tilde{\psi}$ that satisfies: (1) limited liability for the manager ($\tilde{\psi} \leq \tilde{\theta}$); (2) limited liability for the investor ($0 \leq \tilde{\psi}$); and (3) monotonicity in the

⁷ We justify this assumption by noting that there are obvious and well-known restrictions and frictions that make short-selling difficult, and these restrictions have become even more onerous with the SEC’s September 2008 actions (SEC’s “Financial Firm Emergency Order” No. 34-58592; Sept. 18, 2008). Even if we allowed for shorting, the informed trader could make no extra profit by shorting, because any decision to short would reveal her information to the market maker.

⁸ Dye (1998) uses a similar information representation to model privately informed investors.

⁹ The key feature of the information structure is that it is “tiered,” with the manager always weakly better informed than the informed trader, and the informed trader always weakly better informed than the market maker. While we have assumed that—when the informed trader knows that the manager is withholding information—the informed trader does not do anything about what information the manager is withholding, we could have assumed that the informed trader imperfectly knows what the manager knows. We expect that our main results would continue to hold in the latter case, with few changes. We know, for example, that our results are valid with no changes whatsoever if we assume the informed trader learns that, when the manager has withheld information, the informed trader learns into which element of the partition $\{\{\theta_h\}, \{\theta_m, \theta_l\}\}$ the manager’s private information falls.

sense that both $\psi(\theta)$ and $\theta - \psi(\theta)$ are weakly increasing in θ , with $\theta_l - \psi(\theta_l) < \theta_h - \psi(\theta_h)$. In the following, we collectively refer to (1)–(3) as the “limited liability and monotonicity assumptions.” Assumptions similar to these natural, standard assumptions have been adopted in many studies of securities’ design.¹⁰

Preliminary Analysis

We start the analysis by observing that there are only three possible equilibrium levels for the volume, X , of the aggregate order flow for the security. This follows from the assumption that, when the investor experiences a liquidity shock, the investor must sell 100 percent of her security. Since the informed trader has an interest in masking the identity of her trades, either she does not participate in the securities market, or else she submits a market order for exactly the same quantity of the security $\tilde{\psi}$ the investor sells (which we normalize to 1).¹¹ Hence, X ’s only possible values are -1 (if the investor experiences a liquidity shock and the trader stays out of the market), 0 (if either (a) the investor experiences a liquidity shock and the trader trades or (b) neither the investor experiences a liquidity shock nor does the trader trade), or $+1$ (if the investor does not experience a liquidity shock and the trader trades).

Next, we consider how the security $\tilde{\psi}$ is priced by the market maker. Upon observing the manager’s disclosure (if any) and the aggregate order flow X , the market maker posts a price $P(d, X)$ for the security, such that $P(d, X) = E(\tilde{\psi} | d, X)$. This price is what the investor receives by selling her security. On average, an investor who faces a liquidity shock must anticipate having to sell her security for less than its expected value because of the presence of the informed trader.¹² So, the *ex ante* value of the security to the investor:

$$I(\tilde{\psi} = d) \equiv vE(P(d, X) | \tilde{q} = 1) + (1 - v)E(\tilde{\psi}), \quad (1)$$

typically will be below the unconditional expected value $E(\tilde{\psi})$ of the security. The difference between $E(\tilde{\psi})$ and $I(\tilde{\psi}, d)$ is the firm’s “cost of capital” (*CC*) in this model:

$$CC(\tilde{\psi}, d) \equiv E(\tilde{\psi}) - I(\tilde{\psi}, d). \quad (2)$$

The Downside Risk and Upside Potential of Securities

In the following analysis, it is often desirable to characterize a security in terms of how the value of the security changes with the firm’s realized cash flows, as in the following definition.

Definition 1: (a) The “downside risk” of security (ψ_l, ψ_m, ψ_h) is given by $\Delta_D = \psi_m - \psi_l$;
(b) the “upside potential” of security (ψ_l, ψ_m, ψ_h) is given by $\Delta_U = \psi_h - \psi_m$.

The downside risk of a security is the amount by which the payoff of the security drops as a consequence of having the low state θ_l occur, rather than the medium state, and the upside potential of a security is the amount by which the payoff of the security increases as a consequence of having the high state θ_h occur rather than the medium state θ_m . In view of the relationship $\psi_m = \psi_l + \Delta_D$ and $\psi_h = \psi_l + \Delta_D + \Delta_U$, a security with payoffs (ψ_l, ψ_m, ψ_h) can be described

¹⁰ See, e.g., Harris and Raviv (1990), Innes (1990), Nachman and Noe (1994), DeMarzo and Duffie (1999).

¹¹ Were the informed trader to submit market orders in quantities different from that of the investor, she would reveal her trades and thereby eliminate her information advantage over the market maker.

¹² This fact establishes, as was asserted earlier, that were the investor not subject to a liquidity shock, the investor would not participate in this market. Instead, she would wait until the firm’s cash flows realizes its value and collect the cash-flow-contingent payment as specified by the security.

equivalently via the triple $(\psi_l, \Delta_D, \Delta_U)$. It is straightforward to show that, in terms of this triple, the limited liability and monotonicity assumptions are equivalent to the following three conditions: (1) $0 \leq \Delta_i \leq e$ for $i = D, U$, (2) $\Delta_i \neq 0$ for at least one $i = D, U$, (3) $0 \leq \psi_l \leq \theta_l = \mu - e$.

Debt securities, which take the form $\tilde{\psi} = \min\{F, \tilde{\theta}\}$ for some face value F , play an important role in the following. Definition 2 classifies debt instruments in terms of the triple $(\psi_l, \Delta_D, \Delta_U)$.

Definition 2: A security $(\psi_l, \Delta_D, \Delta_U)$ is “risk-free debt” when it is of the form $(\psi_l, 0, 0)$; it is “investment-grade debt” when it is of the form $(\mu - e, \Delta_D, 0)$ where $\Delta_D > 0$; and it is “speculative (or junk) debt” when it is of the form $(\mu - e, e, \Delta_U)$ where $\Delta_U > 0$.

The classification of debt securities makes clear that the basic difference among different types of debt securities is determined by the states in which default occurs (or, alternatively, the probability of default): a firm never defaults on risk-less debt, it defaults on investment-grade debt only if the low state occurs, and it defaults on speculative debt if anything other than the high state occurs.

Equity securities take the form $(\psi_l, \Delta_D, \Delta_U) = (\alpha(\mu - e), \alpha e, \alpha e)$ for some $\alpha \in (0, 1]$; hybrid securities are securities that are neither debt nor equity. Below we show that, when all disclosures are voluntary, both equity securities and hybrid securities are dominated by debt securities, and that equity securities may be optimal when some disclosures are mandatory.

Capital Structure-Dependent Disclosure Policies

Lemma 1 characterizes how the manager’s preferred disclosure policy varies with the firm’s capital structure/security design choice $\tilde{\psi}$.

Lemma 1: Suppose the manager offers the investor a security $(\psi_l, \Delta_D, \Delta_U)$ that satisfies the limited liability and monotonicity conditions. Then, the possible disclosure policies $d(\cdot)$ the security can induce the manager to adopt are the following: $d(\theta_h) = \theta_h$, $d(\theta_l) = \emptyset$, and:

$$\text{if } pe < \Delta_D - (1 - p)\Delta_U, \quad \text{then } d(\theta_m) = \emptyset; \quad (3)$$

$$\text{if } pe > \Delta_D - (1 - p)\Delta_U, \quad \text{then } d(\theta_m) = \theta_m; \quad (4)$$

and if each strict inequality in (3) and (4) is replaced by an equality, then the security can induce the manager either to disclose, or to not disclose, $\tilde{\theta} = \theta_m$.

Lemma 1 shows that any security that satisfies the monotonicity and limited liability conditions always induces the manager to disclose θ_h and never induces him to disclose θ_l , but whether the security induces the manager to disclose θ_m depends upon how large the security’s upside potential is as compared to its downside risk. The proposition thus formalizes one of the claims made in Section III: a firm’s capital structure—that is, the securities it uses to raise capital from investors—affects the firm’s disclosure policy.

The intuition for the proposition is as follows. Consider downside risk first and recall that the manager’s disclosure decision is based on whether the disclosure increases the market’s perceptions of the value of his residual claim $\theta - \psi(\theta)$. It follows that, when the value of his residual claim rises substantially due to having the medium state rather than the low state occur, the manager wants the capital market to know it. Ergo, as the security’s downside risk decreases, the manager is more likely to prefer disclosing θ_m . A similar analysis shows that the manager’s preference for disclosing θ_m also depends on the upside potential of the security. When the inequality in (3) and (4) holds, we say that security $\tilde{\psi}$ induces the manager to adopt a *limited (expansive) disclosure policy* and also that $d(\cdot)$ is the *disclosure policy induced by $\tilde{\psi}$* .

It is also worth noting that Lemma 1 is consistent with, and extends, some classic comparative statics results of voluntary disclosure. For example, the proposition implies that, as e decreases—i.e., as the volatility of the firm's cash flow decreases (in the sense of second-order stochastic dominance)—then a firm is less likely to engage in expansive disclosure. This implication is consistent with Jung and Kwon's (1988) comparative static that, as a firm's cash flow volatility decreases, the threshold above which it discloses information (and below which it withholds information) increases. Also, by rewriting Lemma 1 so as to emphasize how the switch from limited to expansive disclosure depends on the probability p that the firm receives information, we see that there is a threshold probability $p^* \equiv \frac{\Delta_D - \Delta_U}{e - \Delta_U}$ such that the manager engages in expansive (limited) disclosure when $p > p^*$ ($p < p^*$). This is consistent with Dye's (1985) and Jung and Kwon's (1988) result that the disclosure threshold declines—and firms are more likely to disclose their private information—as the probability they receive information increases.

Informed Trader's Optimal Trading Strategy

When the manager discloses the information he received, the market maker is at informational parity with the informed trader. Hence, the trader does not trade, since she knows there is no opportunity to earn positive expected trading profits.¹³ But, recall that, when the manager discloses nothing, the trader, unlike the market maker, knows whether the nondisclosure is due to the manager not receiving information or due to the manager withholding information. The insider will recognize that the market maker is likely to underprice (overprice) the security when the insider knows the reason the manager made no disclosure is because the manager did not receive (withheld) information, and this likely underpricing (overpricing) will trigger the insider to trade (stay out of the market). This provides the intuition for Lemma 2.

Lemma 2: The informed trader's optimal trading strategy is to submit a market order to buy security $\tilde{\psi}$ when the manager makes no disclosure and the trader knows that the reason the manager made no disclosure was because the manager did not receive information; in all other circumstances, the trader does not trade.¹⁴

Determinants of a Firm's Cost of Capital

In this section, we provide a detailed analysis of how a firm's cost of capital depends on the design of its securities and the firm's disclosure policy.

For a risk-free security, that is, a security whose payout $\psi(\theta)$ is constant across all realizations of θ , the cost of capital is 0 because the market maker never misprices such a security. Even a risky security—i.e., one whose payout $\psi(\theta)$ varies with θ —is correctly priced by the market maker in three situations: (1) the manager discloses the firm's realized cash flows; (2) the manager makes no disclosure and the aggregate market order for security $\tilde{\psi}$ is $X = 1$ (because $X = 1$ occurs only when the trader places a market order and the investor did not experience a liquidity shock, which reveals that the trader knew the manager's nondisclosure was due to the manager not receiving information); or (3) the manager makes no disclosure and the aggregate market order for security $\tilde{\psi}$ is $X = -1$ (because $X = -1$ occurs only when the investor, and not the trader, places a market order, which reveals that the trader knew the manager's nondisclosure was due to the manager withholding information).

What creates trading losses for an investor who experiences a liquidity shock is the situation

¹³ We assume that traders who make 0 expected trading profits do not trade. This assumption is justified by any, even trivial, transaction costs from trading.

¹⁴ The proof of Lemma 2 is straightforward and omitted.

in which the manager makes no disclosure and $X = 0$. From the market maker's perspective, this last event is consistent with either: (4) neither the investor nor the informed trader places a market order for security $\tilde{\psi}$, or (5) both the investor (because she faced a liquidity shock) and the market maker (because he knew that the manager's nondisclosure was due to the manager not having received information) place a market order for security $\tilde{\psi}$. In this case, the trader is better informed than the market maker, giving rise to the investor's trading losses. By determining how probable this situation is, and the size of the investor's expected trading losses when it occurs, we can calculate the firm's cost of capital. The next proposition displays the results of this calculation, for any security and any disclosure policy. (In stating Proposition 1, we write $d = 1$ for an expansive disclosure policy (in place of $d(\theta_m) = \theta_m$) and $d = 0$ for a limited disclosure policy (in place of $d(\theta_m) = \emptyset$).

Proposition 1: The firm's cost of capital when it issues security $\tilde{\psi} = (\psi_I, \Delta_D, \Delta_U)$ and adopts disclosure policy d is given by:

$$CC(\tilde{\psi}, d) = A_U^d \times \Delta_U + A_D^d \times \Delta_D, \quad (5)$$

where

$$A_U^d = v(1-p) \frac{\eta}{2} \frac{p(1-v) \left(\frac{\eta}{2} + (1-d)(1-\eta) \right)}{(1-p)v + p(1-v) \left(\frac{\eta}{2} + (1-d)(1-\eta) \right)};$$

$$A_D^d = v(1-p) \frac{\eta}{2} \frac{p(1-v) \left(\frac{\eta}{2} + d(1-\eta) \right)}{(1-p)v + p(1-v) \left(\frac{\eta}{2} + (1-d)(1-\eta) \right)}.$$

Before discussing Proposition 1, we introduce Corollary 1, which helps to illustrate some of the economic consequences of the proposition.

Corollary 1:

- (a) $0 < A_D^0 < A_U^1 < A_U^0 < A_D^1 < \frac{\eta}{2}$; and
- (b) $A_D^0 + A_U^0 < A_D^1 + A_U^1 < \frac{\eta}{2}$.

Corollary 1 shows that the firm's cost of capital varies with each of: (1) the security's downside risk Δ_D and its upside potential Δ_U ; (2) the specifics of the information and trading environments in which the securities are traded (i.e., the probability that the manager receives information p and the probability that the investor is subject to a liquidity shock v); (3) the technology that produces the firm's cash flows (i.e., the probability distribution of $\tilde{\theta}$, as summarized by the parameter η); and (4) the manager's disclosure policy d .

The ordering of the coefficients $A_U^1 < A_U^0$ shows that the upside potential of a security contributes less to the firm's cost of capital for an expansive disclosure policy than for a limited disclosure policy. Stated differently, *more disclosure favors securities with higher potential*. By contrast, the result $A_D^1 > A_D^0$ in Corollary 1 shows that the downside risk of a security contributes more to the firm's cost of capital when the firm's disclosure policy is expansive ($d = 1$) than when the firm's disclosure policy is limited ($d = 0$). Stated alternatively, *less disclosure favors securities with higher downside risk*.

The intuition for the result is that the firm's cost of capital is the same as the investor's expected trading losses. The investor bears trading losses only when she sells her security following a liquidity shock and the informed trader buys the security. Note that the probability these two events both occur is independent of whether the manager adopts an expansive or limited disclosure policy, as this probability is $\nu \times (1 - p)$ in all cases.¹⁵ Hence, the only reason the investor's expected trading losses would be higher under one disclosure policy than another is if the equilibrium price at which the investor sells her security when the aggregate order flow is 0 is lower under one disclosure policy than the other. To isolate the effects of expanded disclosure on a security's downside risk, consider a security that has only downside risk (i.e., $\Delta_U = 0$). Such a security provides two distinct possible payoffs to its holder, a "low" payoff when $\tilde{\psi} = \psi_l$ or a "high" payoff (which is the same whether $\tilde{\psi} = \psi_m$ or $\tilde{\psi} = \psi_h$). Under the expansive disclosure policy, the manager only withholds information when the security's payoff is "low," while under the limited disclosure policy, the manager withholds information both when the security's payoff is "low" and also sometimes when the security's payoff is "high" (specifically, when $\tilde{\psi} = \psi_m$). Consequently, when the market maker calculates the equilibrium price when $d = \emptyset$ and $X = 0$ for a security with no upside potential, the calculation will result in a lower price when the manager adopts the expansive disclosure policy than when he adopts the limited disclosure policy. This provides intuition for the ordering $A_D^1 > A_D^0$.

The result $A_U^1 < A_D^0$ in Corollary 1 shows that the upside potential of a security contributes less to the firm's cost of capital for an expansive disclosure policy than for a limited disclosure policy. Stated differently, *more disclosure favors securities with higher upside potential*. The intuition proceeds along lines similar to that in the preceding paragraph.¹⁶

The results reported in Corollary 1 also show that, if the upside potential and downside risk of a security are the same, then the firm's cost of capital is higher under expansive disclosure than under limited disclosure; that is, with $\Delta \equiv \Delta_D = \Delta_U$, $CC(\tilde{\psi}, 0) = (A_D^0 + A_U^0) \times \Delta < (A_D^1 + A_U^1) \times \Delta = CC(\tilde{\psi}, 1)$.¹⁷

The orderings depicted in the corollary are important to studying both the design of optimal securities while holding a firm's disclosure policy fixed, and the optimal choice among disclosure policies, as will be seen from the results in Propositions 3, 4, and 5 below.¹⁸

¹⁵ This probability $\nu(1 - p)$ follows from the fact that ν is the probability of a liquidity shock, and the informed trader only buys when she knows that the reason for the manager not making a disclosure is that the manager did not receive any information, which occurs with probability $1 - p$.

¹⁶ To isolate the effects of expanded disclosure on a security's upside potential, consider a security that has only upside potential (so $\Delta_D = 0$). Such a security provides two distinct possible payoffs to its holder, a "low" payoff (which is the same whether $\tilde{\psi} = \psi_l$ or $\tilde{\psi} = \psi_m$) or a "high" payoff when $\tilde{\psi} = \psi_h$. Under the limited disclosure policy, the manager always withholds information when the security's payoff is "low" (i.e., regardless of whether $\tilde{\psi} = \psi_l$ or $\tilde{\psi} = \psi_m$), while under the expansive disclosure policy, the manager sometimes discloses information even though the security's payoff is "low" (specifically, when $\tilde{\psi} = \psi_m$). Consequently, when the market maker calculates the equilibrium price when $d = \emptyset$ and $X = 0$ for a security with no downside risk, the calculation will result in a lower price when the manager adopts the limited disclosure policy than when he adopts the expansive disclosure policy. This provides intuition for the ordering $A_U^1 < A_U^0$.

¹⁷ When the aggregate order flow is 0 and the manager makes no disclosure, under expansive disclosure the security's equilibrium price depends on the manager withholding information only when the security's payoff is ψ_l , whereas under the limited disclosure policy the manager withholds information both when the security's payoff is ψ_l and when its payoff is ψ_m . Since, for a security with equal downside risk and upside potential, $\psi_m = E[\tilde{\psi}] > \psi_l$ (as: $E[\tilde{\psi}] = \psi_m$ for a security with equal downside risk and upside potential), it follows that the firm's equilibrium price will be lower when $d = \emptyset$ and $X = 0$ under the expansive disclosure policy than under the limited disclosure policy, which explains why $CC(\tilde{\psi}, 0) < CC(\tilde{\psi}, 1)$ in that case. This provides intuition for the ordering $A_D^0 + A_U^0 < A_D^1 + A_U^1$.

¹⁸ There are many additional comparative statics that can be derived from Proposition 1. To save space, these comparative statics are not presented here, but they are available from the authors.

III. OPTIMAL SECURITY DESIGN AND DISCLOSURE POLICY

In this section, holding fixed the manager's expansive or limited disclosure policy, we identify the manager's most preferred security among all securities that satisfy the limited liability and monotonicity conditions.

It is intuitive, and it is possible to show formally, that an optimal security minimizes the firm's cost of capital, $A_D^d \Delta_D + A_U^d \Delta_U$, subject to a suitable set of constraints. When the security is designed to induce the manager to implement disclosure policy $d \in \{0, 1\}$, the program that identifies the optimal security is the following:

Optimal Security Design d (OSD d) Program:

$$\min_{\psi, \Delta_D, \Delta_U} A_D^d \Delta_D + A_U^d \Delta_U \quad (6)$$

subject to:

$$(1 - 2d)pe \leq (1 - 2d)(\Delta_D - (1 - p)\Delta_U); \quad (7)$$

$$1 \leq I = E\tilde{\psi} - A_D^d \Delta_D - A_U^d \Delta_U; \quad (8)$$

$$0 \leq \psi_i \leq \mu - e; \text{ and} \quad (9)$$

$$0 \leq \Delta_i \leq e, \quad i \in \{\mu, h\}. \quad (10)$$

The OSD d Program incorporates several results derived above: constraint (7) requires that the manager's adopted disclosure policy d be incentive compatible with the manager's security design choice (Lemma 1); constraint (8) requires that the investor be willing to purchase the security (Lemma 2); and constraints (9) and (10) reflect the monotonicity and limited liability constraints. A security is said to be *feasible* if it satisfies all these constraints.

Clearly, a lower bound on the firm's cost of capital is 0. When this bound can be obtained, we say that it is "first-best." The first-best situation arises if and only if $\Delta_D = \Delta_U = 0$, i.e., the security issued is risk-free debt. The maximum amount of capital the manager can raise by issuing a risk-free security is $\mu - e$. Hence, first-best financing is achieved if and only if the required capital to implement the investment ($= 1$) is less than $\mu - e$, i.e., if and only if the volatility e in the firm's cash flows does not exceed $\mu - 1$.

The next proposition reports that the risk-free debt security achieves first-best and induces the manager to adopt the expansive disclosure policy.

Proposition 2: There is a security that implements the expansive disclosure policy that yields the first-best outcome if and only if $e \leq \mu - 1$. That security must take the form of risk-free debt with face value $F \in [1, \mu - e]$; it always induces expansive disclosure.

Proposition 2 provides the first illustration of the sensitivity of the structure of optimal securities to the size of the firm's cash flow volatility, as measured by the parameter e . This sensitivity will be shown to be a central feature of the analysis that follows. Since Proposition 2 establishes the optimal security for $e \leq \mu - 1$, the following analysis focuses on the case of $e > \mu - 1$.

Optimal Securities that Implement Limited Disclosure

This section studies optimal securities that implement the limited disclosure policy. The main result is to establish the existence of a hierarchy of optimal securities, where what constitutes an optimal security varies with the volatility parameter e .

To proceed, we observe that we can set $\psi_l = \mu - e$ in Program OSD0 with no loss in profits to the manager. With this substitution, it is clear that Program OSD0 is a linear program in the two variables Δ_D and Δ_U that can be solved geometrically. To obtain the optimal security, all that has to be done is to search for the lowest iso-cost line for the objective function, i.e., the lowest value k among the lines $A_D^0 \Delta_D + A_U^0 \Delta_U = k$ that lies within this constraint set. Figure 2 illustrates the linear program that yields the optimal security that implements limited disclosure.

The following proposition summarizes the findings from the geometric investigation of the OSD0 Program in the Appendix.

Proposition 3: The optimal security that implements the limited disclosure policy is:

- (1) investment-grade debt with face value:

$$\mu - (1-p)e \quad \text{if } \mu - 1 < e \leq \frac{\mu - 1}{1 - \left(1 - \frac{\eta}{2} - A_D^0\right)p} \equiv \bar{e}_{ID}^0;$$

- (2) investment-grade debt with face value:

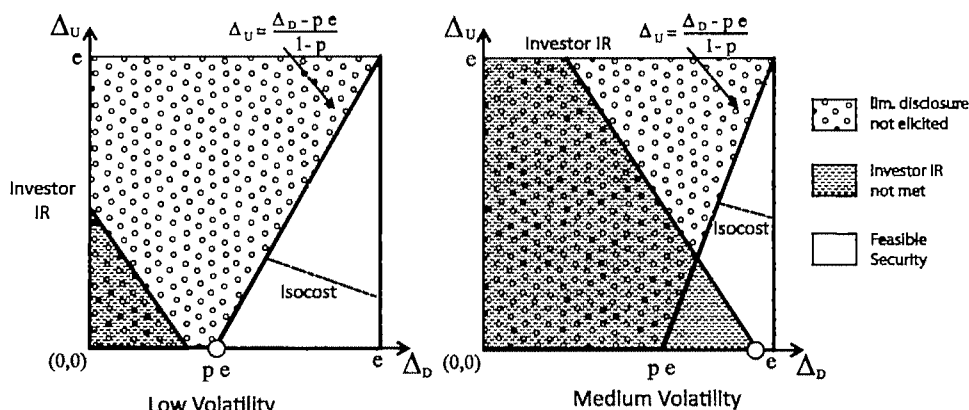
$$\mu - e + \frac{1 + e - \mu}{1 - \frac{\eta}{2} - A_D^0} \quad \text{if } \bar{e}_{ID}^0 < e \leq \frac{\mu - 1}{\frac{\eta}{2} + A_D^0} \equiv \bar{e}_{SD}^0;$$

- (3) speculative debt with face value:

$$\frac{\mu \left(\frac{\eta}{2} - A_U^0 - 1 \right) + 1 + \left(\frac{\eta}{2} + A_D^0 \right) e}{\frac{\eta}{2} - A_U^0} \quad \text{if } \bar{e}_{SD}^0 < e < \frac{\mu - 1}{A_D^0 + A_U^0} \equiv \bar{e}_{SD}^0; \quad \text{and}$$

- (4) the limited disclosure policy cannot be implemented by any feasible security if $e \geq \bar{e}_{SD}^0$.

FIGURE 2
Optimal Securities that Implement the Limited Disclosure Policy



Proposition 3 establishes that there is a simple two-level hierarchy of optimal securities that induce the manager to implement the limited disclosure policy: investment-grade debt is optimal when the volatility of the firm's cash flows is sufficiently low, and speculative debt is optimal when the volatility of the firm's cash flows is high. When the firm's cash flow volatility is extremely high, no feasible security of any kind implements the limited disclosure policy.

The intuition for the results in Proposition 3 is clear: as noted in the discussion of Proposition 2 above, risk-free debt never induces the manager to implement the limited disclosure policy. Among the risky securities that induce the manager to implement the limited disclosure policy, the manager prefers those securities whose payoffs do not vary much with the firm's cash flows because the absence of much volatility in a security's payoff reduces the informational advantage of the informed trader over the market maker, thereby reducing the trader's expected trading profits and driving down the firm's cost of capital. More than this can be said, however. When choosing among those risky securities that induce the manager to adopt the limited disclosure policy, the manager prefers securities that have only downside risk. Such securities have a lower cost of capital than do securities that also have upside potential, for the reasons described in the discussion following Corollary 1.

When $\psi_l = \mu - e$, securities with only downside risk are of the form $(\psi_l, \Delta_D, \Delta_U) = (\mu - e, \Delta_D, 0)$; such securities are investment-grade debt. An optimal security that implements the limited disclosure policy is, therefore, investment grade debt with downside risk Δ_D that satisfies the disclosure constraint in Equation (7). This explains part (1) of Proposition 3.

As the volatility in the firm's cash flows increases, the volatility of the payoff of the optimal security must increase as well. The cost of issuing a security with more volatile payoffs is that the information advantage of the trader over the market maker goes up, which increases the investor's expected trading losses. In order to compensate the investor for her increased trading losses, the manager must issue debt with a higher face value, resulting in a higher cost of capital. This explains part (2) of Proposition 3.

For sufficiently high cash flow volatility levels, even securities with maximal downside risk (but no upside potential) are unable to induce the investor to supply financing. To construct feasible securities in such cases, the manager must resort to securities that have both downside risk and upside potential. Such securities divert all the firm's cash flows to the security holder, unless the most favorable (highest cash flow) state occurs. These securities are speculative debt. This explains part (3) of Proposition 3. Financing the project eventually becomes infeasible at very high levels of the firm's cash flow volatility, because the investor's expected trading losses become so high that no feasible security would compensate her adequately for her expected trading losses. This explains part (4) of Proposition 3.

Optimal Securities that Implement Expansive Disclosure

The program that determines the optimal security that implements the expansive disclosure policy is also a linear program in the two variables Δ_D and Δ_U , which can be solved geometrically. As was the case for the linear program in the preceding section, to obtain the optimal security, all that has to be done is to search for the lowest iso-cost line for the objective function, i.e., the lowest value k among the lines $A_D^1 \Delta_D + A_U^1 \Delta_U = k$ that lies within this constraint set. Figure 3 illustrates the linear program that yields the optimal security that implements expansive disclosure.

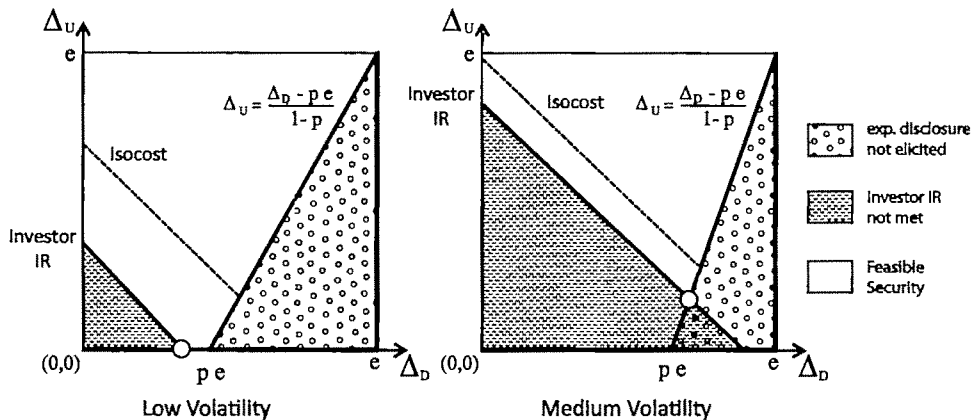
The following proposition summarizes the findings from this geometric investigation.

Proposition 4: An optimal security that implements the expansive disclosure policy is:

- (1) risk-free debt with face value:

$$F \in [1, \mu - e] \quad \text{if } e \leq \mu - 1;$$

FIGURE 3
Optimal Securities that Implement the Expansive Disclosure Policy



- (2) investment-grade debt with face value:

$$e + \frac{1 + e - \mu}{\frac{\eta}{2} - A_U^1} \quad \text{if } \mu - 1 < e \leq \frac{\mu - 1}{1 - \left(1 - \frac{\eta}{2} - A_D^1\right)p} \equiv \bar{e}_{ID}^1;$$

- (3) a hybrid security if:

$$\bar{e}_{ID}^1 < e < \frac{\mu - 1}{A_D^1 + A_U^1} \equiv \bar{e}_{ID}; \text{ and}$$

- (4) the expansive disclosure policy cannot be implemented by any feasible security if $e \geq \bar{e}_{ID}^1$.

Proposition 4 establishes that there is a simple three-level hierarchy of optimal securities that induce the manager to implement the expansive disclosure policy: risk-free debt is optimal when the volatility of the firm's cash flows is extremely low; investment-grade debt is optimal when the volatility of the firm's cash flows is too high to make risk-free debt feasible, but is still sufficiently low; and hybrid securities are optimal for higher levels of volatility. As was also true for the limited disclosure policy discussed in the previous section, when the firm's cash flow volatility is extremely high, no feasible security of any kind implements the expansive disclosure policy.

The intuition for part (1) was explained in the discussion of Proposition 2. The intuition for parts (2) and (3) is somewhat similar to that for Proposition 3: as the volatility in the firm's cash flows increases, the volatility of the feasible securities that induce the investor to supply financing also increases. In the discussion following Corollary 1, we observed that the investor's trading losses depend on the price the market maker sets following no disclosure and an order flow of 0. When the manager implements the expansive disclosure policy, the non-disclosure price is a weighted average of ψ_l (in case the manager strategically withheld information) and $E[\tilde{\psi}] = \frac{\eta}{2}\psi_l + (1 - \eta)\psi_m + \frac{\eta}{2}\psi_h$ (in case the manager did not receive information). Hence, the non-disclosure price overweights only the low outcome of the security (relative to the prior probability). In order to minimize the investor's trading losses—or, equivalently, minimize the firm's cost of capital—

the manager chooses a security that sets the payoff in the low state to its maximum value, $\psi_l = \mu - e$, but is indifferent between increasing the security's payoff in the intermediate or high state. It follows that debt securities are the optimal securities that induce the expansive disclosure policy (to the extent they satisfy the monotonicity and limited liability constraints). To compensate the investor for her trading losses as the volatility in the firm's cash flows increases, the manager must offer a higher face value on its debt instrument.

However, doing so also reduces the manager's incentives to implement the expansive disclosure policy. When the cash flow volatility extends beyond the critical threshold e_{ID}^1 , no feasible investment-grade debt will continue to induce the manager to adopt the expansive disclosure policy. For higher volatilities, the manager needs to issue hybrid securities with not only downside risk, but also upside potential, in order to induce the manager to disclose the intermediate outcome (as required under expansive disclosure policies). Part (4) of Proposition 4 follows for reasons similar to part (5) of Proposition 3: at very high levels of cash flow volatility, the manager cannot finance the project.

Global Hierarchy of Securities and Disclosure Policies

So far, we have established the existence of two hierarchies of securities, one for securities that optimally implement the limited disclosure policy and one for securities that optimally implement the expansive disclosure policy. We next develop a combined hierarchy that involves optimizing across both securities and disclosure policies. We call the optimal security and disclosure policy that solves this joint optimization problem the "globally optimal" security and disclosure policy.

By relying on the results already established, we can get an immediate sense of what constitutes the globally optimal security and disclosure policy in the "corner" case in which the firm's cash flow volatility is sufficiently low (i.e., $e \leq \mu - 1$). For that case, Proposition 2 establishes that the expansive disclosure policy necessarily dominates because risk-free securities induce the expansive disclosure policy but not the limited disclosure policy. Appealing to the continuity of our optimization program, the expansive disclosure policy is also the globally optimal disclosure policy for cash flow volatilities e close to but higher than $\mu - 1$. In that case, the associated globally optimal security is investment-grade debt (by Proposition 4, part 2).

We can also get a sense for what constitutes the globally optimal security and disclosure policy in the other "corner" case in which the firm's cash flow volatility is sufficiently high. When the firm's cash flow volatility is sufficiently high, the only potentially feasible security is one that transfers nearly all of the firm's cash flows to the investor—i.e., one for which both the downside risk and upside potential of the security are nearly maximal: $\Delta_D \approx e \approx \Delta_U$. Given this observation, then the result in Corollary 1 that $A_D^0 + A_U^0 < A_D^1 + A_U^1$ leads immediately to the conclusion that the globally optimal disclosure policy must be the limited disclosure policy, as that result implies that a firm's cost of capital under the limited disclosure policy, $A_D^0 \Delta_D + A_U^0 \Delta_U \approx (A_D^0 + A_U^0)e$, will be less than its cost of capital under the expansive disclosure policy, $A_D^1 \Delta_D + A_U^1 \Delta_U \approx (A_D^1 + A_U^1)e$. Appealing to continuity a second time, we can further conclude that the limited disclosure policy is the globally optimal disclosure policy for all cash flow volatilities close to but lower than the maximum volatility for which the project can be implemented. The associated globally optimal security is speculative debt (by Proposition 3).

Having settled the question of what constitutes the globally optimal security and disclosure policy when the firm's cash flow volatility falls in the "corners," it remains to determine the globally optimal security and disclosure policy when the cash flow volatility assumes "intermediate" values. The next proposition provides a complete answer to this question.

Proposition 5: There exists a unique value, e^* , for the volatility of the firm's cash flows, with $e^* \in (\mu - 1, \underline{e}_{ID}^1)$, such that:

- (1) if $0 < e \leq e^*$, then expansive disclosure combined with either riskless debt (when $e \leq \mu - 1$) or investment-grade debt (when $e > \mu - 1$), as described in Proposition 4, is optimal;
- (2) if $e^* < e < \bar{e}_{SD}^0$, then limited disclosure combined with either investment-grade debt (when $e \leq \underline{e}_{SD}^0$) or speculative debt (when $e > \underline{e}_{SD}^0$), as described in Proposition 3, is optimal;¹⁹ and
- (3) if $e \geq \bar{e}_{SD}^0$, then the project cannot be implemented by any feasible security.

Proposition 5 establishes that the manager prefers the expansive disclosure policy in all environments characterized by a sufficiently low level of cash flow volatility, whereas he prefers the limited disclosure policy in all environments characterized by higher levels of cash flow volatility, up to some maximum cash flow volatility beyond which the project cannot be implemented by any disclosure policy (either limited or expansive). In addition, Proposition 5 establishes that, for all cash flow volatilities for which the manager can implement the project, the optimal security is always a debt security. The debt security that is globally optimal is riskless debt when the cash flow volatility is sufficiently low, becoming investment-grade debt as the cash flow volatility increases, and speculative debt as the cash flow volatility increases still further.

Another feature of Proposition 5 worth mentioning is that hybrid securities are never optimal. Hybrid securities have one function: to induce expansive disclosure for parameter values such that standard debt alone would not induce expansive disclosure. Relative to the standard debt security, the hybrid security exhibits more upside potential and less downside risk. Therefore, the manager trades off limited-disclosure downside risk (when using standard debt) and expansive-disclosure upside potential (when using the hybrid). However, following the intuition given in Corollary 1, the limited-disclosure downside risk is less costly than the expansive-disclosure upside potential. It then follows that using the standard debt instrument is preferable to encouraging expansive disclosure using the hybrid security.

Perhaps the most important prediction that emerges from Proposition 5 is the following corollary:

Corollary 2: In equilibrium, there is a negative association between firms' cost of capital and the amount of information voluntarily disclosed by firms.

This prediction follows from two previously established facts. First, as a firm's cash flow volatility e increases, its maximum expected profits decline due to the firm's cost of capital increasing (this follows by combining Propositions 3, 4, and 5). Second, as a firm's cash flow volatility increases, it optimally switches from the expansive disclosure policy to the limited disclosure policy (by Proposition 5).

This negative correlation does not imply, however, that more expansive voluntary disclosure causes firms' cost of capital to decline. On the contrary, the model predicts that, whenever the firm issues risky debt, more disclosure leads to a strict increase in the firm's cost of capital. To see this, first consider investment-grade debt. The cost of capital associated with issuing investment-grade debt, $(\mu - e, \Delta_D, 0)$, is $A_D^q \Delta_D$. Expansive disclosure yields a greater cost of capital for investment-grade debt than does limited disclosure because, by Corollary 1, $A_D^0 < A_D^1$. Next consider speculative debt. The cost of capital associated with issuing speculative debt and adopting the limited disclosure policy is $A_D^0 e + A_D^0 \Delta_U$. This is always less than the cost of capital associated with

¹⁹ The proof of the proposition establishes that $\underline{e}_{ID}^1 < \underline{e}_{ID}^0$. So, since $\underline{e}_{ID}^0 < \underline{e}_{SD}^0$ by Proposition 3, each of the intervals $(e^*, \underline{e}_{SD}^0]$ and $(\underline{e}_{SD}^0, \bar{e}_{SD}^0]$ is nonempty.

issuing speculative debt and adopting the expansive disclosure policy, $A_D^1 e + A_U^1 \Delta_U$ because the downside risk, Δ_D , is higher than the upside potential, Δ_U . Thus, even though for any given risky debt security, there is a positive association between a firm's cost of capital and how much it discloses, when the firm chooses its securities optimally, the equilibrium (hence, predicted) association between a firm's cost of capital and how much the firm discloses is negative.

IV. INTRODUCING MANDATORY DISCLOSURES

In this section, we extend the model by considering how mandatory disclosure requirements affect the manager's choice among securities, the manager's propensity to make supplementary voluntary disclosures, and the firm's cost of capital. We consider two distinct forms of mandatory disclosure requirements. One form, which we label "loss recognition," requires that the manager disclose the low outcome (θ_l) when he knows it. The other form, which we label "gains recognition," requires the manager to disclose the medium outcome (θ_m) when he knows it.

Optimal Security Design under Mandatory Loss Recognition

The next proposition describes the manager's preferred voluntary disclosure policy under mandatory loss recognition. Under mandatory loss recognition, the manager always discloses both the low and the high outcomes. Whether he discloses the medium outcome depends on the relative magnitude of the investor's downside risk and upside potential, as the following proposition indicates:

Proposition 6: Under mandatory loss recognition disclosures, if the manager offers the investor a security $(\psi_l, \Delta_D, \Delta_U)$ that satisfies the limited liability and monotonicity conditions, then the possible voluntary disclosure policies $d(\cdot)$ the security can induce the manager to adopt are the following: $d(\theta_l) = \theta_l$, $d(\theta_h) = \theta_h$, and:

$$\text{if } \Delta_D > \Delta_U, \quad \text{then } d(\theta_m) = \emptyset; \quad (11)$$

$$\text{if } \Delta_D < \Delta_U, \quad \text{then } d(\theta_m) = \theta_m; \quad (12)$$

and if each strict inequality in (11) and (12) is replaced by an equality, then the security can induce the manager either to disclose, or not to disclose, $\tilde{\theta} = \theta_m$.

The intuition for this result is as follows. Whether the manager discloses the medium outcome depends on how the value of his residual claim compares when he discloses θ_m versus when he does not make a disclosure. Since the manager discloses both the low and high outcome under mandatory loss recognition, the value of his residual claim given the medium outcome is higher (lower) when the manager is exposed to more (less) downside risk than upside potential. Equivalently, the manager chooses to disclose (withhold) the medium outcome when the investor captures less (more) of the downside risk than the upside potential, i.e., $\Delta_D < \Delta_U$ ($\Delta_D > \Delta_U$).

Juxtaposing this disclosure policy with Lemma 1 provides a dramatic illustration of how mandatory disclosures can affect a firm's voluntary disclosure policy. Because the inequality (3) characterizing nondisclosure in Lemma 1 can be written alternatively as $\Delta_U < \frac{\Delta_D - pe}{1-p}$, and since, for all $\Delta_D \in [0, e]$, it is clear that $\frac{\Delta_D - pe}{1-p} < \Delta_D$, it follows that any security $(\psi_l, \Delta_D, \Delta_U)$ that induces the manager not to disclose θ_m when there are no disclosure requirements will also induce the manager not to disclose θ_m when the mandatory loss recognition disclosure requirements are imposed. Thus, in the case of mandatory loss recognition disclosures, holding the security the firm issues fixed, mandatory disclosure inhibits voluntary disclosure. We show in Proposition 7 below, however, that if the manager is allowed to optimize his choice among securities (when the dis-

closure regime changes from all disclosures being voluntary to the mandatory loss recognition regime), then mandatory disclosure leads to expanded voluntary disclosure. This conclusion thus provides yet another illustration of the importance of endogenizing a firm's capital structure when making predictions about disclosures: the predictions about disclosures that emerge from holding capital structure fixed are the opposite of the predictions that emerge when the firm's capital structure is endogenized.

When $\Delta_D < \Delta_U$, the manager engages in full disclosure of his private information and the market maker and the investor are at informational parity with the informed trader. Accordingly, the security will always be correctly priced, the investor will not incur any expected trading losses, the firm's cost of capital will be 0, and the first-best outcome will be achieved. The next proposition shows that the manager can exploit this result and achieve first-best by issuing an equity security.

Proposition 7: Under mandatory loss recognition disclosures, for all $e \in (0, \mu)$, the equity security $\tilde{\psi} = \frac{\bar{\theta}}{\mu}$ (i.e., $(\psi_l, \Delta_D, \Delta_U) = (1 - \frac{e}{\mu}, \frac{e}{\mu}, \frac{e}{\mu})$) induces the manager to engage in full disclosure and achieves the first-best outcome.²⁰

Notice that no risky debt security will induce the manager to fully disclose his private information, because no risky debt security satisfies inequality (12). Thus, no risky debt security induces the manager to engage in a policy of full disclosure; hence, no risky debt security completely eliminates the informed trader's information advantage over the market maker and yields the first-best outcome.

We conclude this subsection with two further comments concerning Proposition 7. First, the proposition illustrates that what are considered to be optimal securities can change with the introduction of mandatory disclosure requirements, because equity securities are typically not optimal when all disclosures are voluntary. Second, the proposition illustrates how mandatory disclosures can lead to an increase in the manager's maximum expected profits. We shall see in the next subsection, however, that mandatory disclosures do not always increase the manager's maximum expected profits.

Optimal Security Design under Mandatory Gains Recognition

In this subsection, we discuss the mandatory disclosure regime referred to as "gains recognition" above, which entails that the manager must disclose the medium outcome (θ_m) when he knows it.

Under gains recognition, the manager will always keep silent when he learns that $\tilde{\theta} = \theta_l$ occurs—as he cannot possibly benefit by revealing that the firm's cash flows are the lowest possible level. Because the manager will always disclose $\tilde{\theta} = \theta_h$ when he learns it, and he necessarily discloses $\tilde{\theta} = \theta_m$ when he learns it to be in conformity with the gains recognition mandatory disclosure requirement, it follows that the disclosure policy the manager adopts in the presence of mandatory gains recognition is tantamount to the manager implementing the expansive disclosure policy. From this observation, it is immediate that mandatory gains recognition disclosures sometimes reduces the manager's maximum expected profits relative to the situation in which all disclosures are voluntary. These remarks prove:

Proposition 8: Under mandatory gains recognition disclosures, the disclosure policy the manager is induced to adopt is always the expansive disclosure policy, and

²⁰ The proof of Proposition is immediate: since $\Delta_D = \Delta_U$, condition (12) holds (weakly) for the equity security $(\psi_l, \Delta_D, \Delta_U) = (1 - \frac{e}{\mu}, \frac{e}{\mu}, \frac{e}{\mu})$, so the manager (weakly) prefers full disclosure.

the manager's maximum expected profits are always weakly (in fact, strictly for sufficiently high levels of cash flow volatility) below what the manager could achieve when all disclosures are voluntary.

Viewing mandatory disclosures as a commitment to a disclosure policy, this result shows that commitment to a disclosure policy can reduce the manager's maximum expected profits. While it is commonplace to observe that mandatory disclosure requirements may be undesirable because they generate proprietary costs of disclosure, Propositions 7 and 8 combined indicate that unqualified statements about the desirability or undesirability of expanded mandatory disclosure requirements, or unqualified statements about the desirability of committing to expanded disclosures, typically are invalid even absent concerns about disclosing proprietary information.

V. CONCLUSION

The owners of a firm use all of the instruments at their disposal, including their operating choices, financing choices, investment choices, and disclosure choices, to maximize the expected value of their residual stakes in their firm. Because the array of such instruments available to a firm's owners is in practice so large, any model must inevitably put some restrictions on the range of endogenous variables the owners of a firm are posited to select. The present model demonstrates that, for purposes of studying a firm's disclosure policy, it is undesirable to take the firm's financing decisions as given, because as the firm's financing decisions change, the residual stake over which the owners of the firm have claim change, and the latter change alters the owners' preferred voluntary disclosure policy. We show that the reverse is also true. Namely, the capital structure the owners of a firm prefer also depends on the firm's disclosure policy. These interdependencies imply that, in equilibrium, a firm's capital structure and disclosure policy are jointly determined, and together determine the firm's cost of capital.

APPENDIX

Proof of Lemma 1

In view of the assumed monotonicity of $\theta - \psi(\theta)$, the unique optimality of $d(\theta_h) = \theta_h$ and $d(\theta_l) = \emptyset$ is clear. Consequently, the proof focuses on the specification of $d(\theta_m)$. If Equation (3) describes the manager's disclosure policy, then $E[\tilde{\theta} - \psi(\tilde{\theta}) | d]$ yields:

$$E[\tilde{\theta} - \psi(\tilde{\theta}) | d = \theta_h] = \theta_h - \psi_h$$

and, by Bayes' rule:

$$\begin{aligned} E[\tilde{\theta} - \psi(\tilde{\theta}) | d = \emptyset] &= \frac{(1-p)(\theta_m - E[\tilde{\psi}]) + p\left((1-\eta)(\theta_m - \psi_m) + \frac{\eta}{2}(\theta_l - \psi_l)\right)}{1-p + p\left(1 - \frac{\eta}{2}\right)} \\ &= \theta_m - E[\tilde{\psi}] - \frac{p\left((1-\eta)(\psi_m - E[\tilde{\psi}]) + \frac{\eta}{2}(e - E[\tilde{\psi}] + \psi_l)\right)}{1-p - \frac{\eta}{2}} \end{aligned}$$

where:

$$E[\tilde{\psi}] = (1 - \eta)\psi_m + \frac{\eta}{2}(\psi_l + \psi_h) = \psi_m + \frac{\eta}{2}[(\psi_h - \psi_m) - (\psi_m - \psi_l)].$$

Given the restriction to truth-telling, were the manager to disclose $d = \theta_m$, then necessarily:

$$E[\tilde{\theta} - \psi(\tilde{\theta})|d = \theta_m] = \theta_m - \psi_m.$$

So, for the disclosure policy in Equation (3) to be an equilibrium, we require:

$$E[\tilde{\theta} - \psi(\tilde{\theta})|d = \theta_m] \leq E[\tilde{\theta} - \psi(\tilde{\theta})|d = \emptyset],$$

i.e.:

$$\theta_m - \psi_m \leq \mu - E[\tilde{\psi}] - \frac{p \left((1 - \eta)(\psi_m - E[\tilde{\psi}]) + \frac{\eta}{2}(e - E[\tilde{\psi}] + \psi_l) \right)}{1 - p \frac{\eta}{2}}$$

which simplifies to:

$$pe \leq \Delta_D - (1 - p)\Delta_U.$$

The proof of Equation (4) is analogous and omitted.

Proof of Proposition 1

The investor's expected payoff is:

$$\begin{aligned} I &= (1 - \nu)E[\tilde{\psi}] + \nu \left(\begin{aligned} &\Pr(d = \emptyset, X = -1|q = 1)P(\emptyset, -1) \\ &+ \Pr(d = \emptyset, X = 0|q = 1)P(\emptyset, 0) \\ &+ \Pr(d = \theta_h)\psi_h + \Pr(d = \mu)\psi_m \end{aligned} \right) \\ &= E[\tilde{\psi}] + \nu \left(\begin{aligned} &\Pr(d = \emptyset, X = -1|q = 1)P(\emptyset, -1) \\ &+ \Pr(d = \emptyset, X = 0|q = 1)P(\emptyset, 0) \\ &+ \Pr(d = \theta_h)\psi_h + \Pr(d = \mu)\psi_m - E[\tilde{\psi}] \end{aligned} \right) \end{aligned}$$

Note that:

$$\Pr(d = \emptyset, X = -1|q = 1) + \Pr(d = \emptyset, X = 0|q = 1) + \Pr(d = \theta_h) + \Pr(d = \mu) = 1.$$

The price $P(\emptyset, X)$ can be written as:

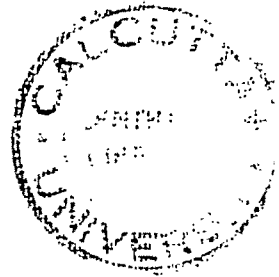
$$\begin{aligned} P(\emptyset, X) &= \psi_m + \Pr(\theta_h|\emptyset, X, d)\Delta_U - \Pr(\theta_l|\emptyset, X, d)\Delta_D \\ &= E[\tilde{\psi}] - \left(\psi_m + \frac{\eta}{2}[\Delta_U - \Delta_D] \right) + \psi_m + \Pr(\theta_h|\emptyset, X, d)\Delta_U - \Pr(\theta_l|\emptyset, X, d)\Delta_D \\ &= E[\tilde{\psi}] + \left(\Pr(\theta_h|\emptyset, X, d) - \frac{\eta}{2} \right) \Delta_U - \left(\Pr(\theta_l|\emptyset, X, d) - \frac{\eta}{2} \right) \Delta_D. \end{aligned}$$

Then, I can be written as:

$$\begin{aligned}
I = E[\tilde{\psi}] + v & \left(\begin{aligned} & \Pr(d = \emptyset, X = -1 | q = 1, d) \left(\begin{aligned} & E[\tilde{\psi}] + \left(\Pr(\theta_h | \emptyset, X = -1, d) - \frac{\eta}{2} \right) \Delta_U \\ & - \left(\Pr(\theta_l | \emptyset, X = -1, d) - \frac{\eta}{2} \right) \Delta_D \end{aligned} \right) \\ & + \Pr(d = \emptyset, X = 0 | q = 1, d) \left(\begin{aligned} & E[\tilde{\psi}] + \left(\Pr(\theta_h | \emptyset, X = 0, d) - \frac{\eta}{2} \right) \Delta_U \\ & - \left(\Pr(\theta_l | \emptyset, X = 0, d) - \frac{\eta}{2} \right) \Delta_D \end{aligned} \right) \\ & + \Pr(d = \theta_h) \left(E[\tilde{\psi}] + \left(1 - \frac{\eta}{2} \right) \Delta_U + \frac{\eta}{2} \Delta_D \right) \\ & + \Pr(d = \mu) \left(E[\tilde{\psi}] - \frac{\eta}{2} (\Delta_U - \Delta_D) \right) - E[\tilde{\psi}] \end{aligned} \right) = E[\tilde{\psi}] \\ + v & \left(\begin{aligned} & \Pr(d = \emptyset, X = -1 | q = 1, d) \left(\begin{aligned} & \left(\Pr(\theta_h | \emptyset, X = -1, d) - \frac{\eta}{2} \right) \Delta_U \\ & - \left(\Pr(\theta_l | \emptyset, X = -1, d) - \frac{\eta}{2} \right) \Delta_D \end{aligned} \right) \\ & + \Pr(d = \emptyset, X = 0 | q = 1, d) \left(\begin{aligned} & \left(\Pr(\theta_h | \emptyset, X = 0, d) - \frac{\eta}{2} \right) \Delta_U \\ & - \left(\Pr(\theta_l | \emptyset, X = 0, d) - \frac{\eta}{2} \right) \Delta_D \end{aligned} \right) \\ & + \Pr(d = \theta_h) \left(1 - \frac{\eta}{2} \right) \Delta_U + \frac{\eta}{2} \Delta_D \\ & - \Pr(d = \mu) \frac{\eta}{2} (\Delta_U - \Delta_D) \end{aligned} \right)
\end{aligned}$$

and therefore:

$$\begin{aligned}
A_U^d &= v \left(\begin{aligned} & \Pr(d = \emptyset, X = -1 | q = 1, d) \left(\frac{\eta}{2} - \Pr(\theta_h | \emptyset, X = -1, d) \right) \\ & + \Pr(d = \emptyset, X = 0 | q = 1, d) \left(\frac{\eta}{2} - \Pr(\theta_h | \emptyset, X = 0, d) \right) \\ & - \left(1 - \frac{\eta}{2} \right) \Pr(d = \theta_h) + \frac{\eta}{2} \Pr(d = \mu) \end{aligned} \right) \\ A_D^d &= v \left(\begin{aligned} & \Pr(d = \emptyset, X = -1 | q = 1, d) \left(\Pr(\theta_l | \emptyset, X = -1, d) - \frac{\eta}{2} \right) \\ & + \Pr(d = \emptyset, X = 0 | q = 1, d) \left(\Pr(\theta_l | \emptyset, X = 0, d) - \frac{\eta}{2} \right) \\ & - \frac{\eta}{2} \Pr(d = \theta_h) - \frac{\eta}{2} \Pr(d = \mu) \end{aligned} \right).
\end{aligned}$$



Market Maker's Beliefs if the Manager Implements the Limited Disclosure Policy

Let $s_{\text{knowcash}} = \theta$ indicate that the manager learns that cash flows will be θ , and $s_{\text{knowcash}} = \emptyset$ indicate that the manager receives no information about cash flows. Suppose the manager does not disclose the median value and the market maker conjectures that the trader does not buy the security if she sees $s_{\text{withheld}} = 1$ and purchases the security if she sees $s_{\text{withheld}} = 0$. Then, the market maker's beliefs are as follows.

If $X = 1$ and $d = \emptyset$, then the trader must have purchased the security because $s_{\text{withheld}} = 0$. Hence, the market maker beliefs equal the prior beliefs.

$$\Pr(\theta_h | \emptyset, X = 1, d = 0) = \Pr(\theta_h | s_{\text{withheld}} = 0) = \frac{\eta}{2}$$

$$\Pr(\theta_l | \emptyset, X = 1, d = 0) = \Pr(\theta_l | s_{\text{withheld}} = 0) = \frac{\eta}{2}.$$

If $X = 0$ and $d = \emptyset$, then either (1) the trader purchased the security because $s_{\text{withheld}} = 0$ and there was a liquidity shock (which occurs with probability $(1 - p)v$), or (2) the trader did not purchase the security because $s_{\text{withheld}} = 1$ (which implies $s_{\text{knowcash}} \in \{\mu, \theta_l\}$) and there was no liquidity shock (which occurs with probability $p(1 - \frac{\eta}{2})(1 - v)$). Hence, the market maker's beliefs are:

$$\Pr(\theta_h | \emptyset, X = 0, d = 0) = \frac{\Pr(\theta_h, d = \emptyset, X = 0)}{\Pr(d = \emptyset, X = 0)} = \frac{\frac{\eta}{2}(1 - p)v}{(1 - p)v + \left(1 - \frac{\eta}{2}\right)p(1 - v)} < \frac{\eta}{2}$$

$$\Pr(\theta_l | \emptyset, X = 0, d = 0) = \frac{\Pr(\theta_l, d = \emptyset, X = 0)}{\Pr(d = \emptyset, X = 0)} = \frac{\frac{\eta}{2}((1 - p)v + p(1 - v))}{(1 - p)v + \left(1 - \frac{\eta}{2}\right)p(1 - v)} > \frac{\eta}{2}.$$

If $X = -1$ and $d = \emptyset$, then from $X = -1$, the market maker can infer that the trader did not purchase the security because $s_{\text{withheld}} = 1$. Hence:

$$\Pr(\theta_h | \emptyset, X = -1, d = 0) = 0$$

$$\Pr(\theta_l | \emptyset, X = -1, d = 0) = \frac{\frac{\eta}{2}}{1 - \eta + \frac{\eta}{2}} = \frac{\eta}{2 - \eta}.$$

Based on these beliefs and:

$$\Pr(d = \emptyset, X = -1 | q = 1, d = 0) = \left(1 - \frac{\eta}{2}\right)p$$

$$\Pr(d = \emptyset, X = 0 | q = 1, d = 0) = (1 - p)$$

$$\Pr(d = \theta_h) = \frac{\eta}{2}p$$

$$\Pr(d = \mu) = 0$$

we can compute:

$$\begin{aligned} A_U^0 &= v \left(\Pr(d = \emptyset, X = -1 | q = 1, d = 0) \left(\frac{\eta}{2} - \Pr(\theta_h | \emptyset, X = -1, d = 0) \right) \right. \\ &\quad \left. + \Pr(d = \emptyset, X = 0 | q = 1, d = 0) \left(\frac{\eta}{2} - \Pr(\theta_h | \emptyset, X = 0, d = 0) \right) \right. \\ &\quad \left. - \left(1 - \frac{\eta}{2}\right) \Pr(d = \theta_h) + \frac{\eta}{2} \Pr(d = \mu) \right) \\ &= v \left(\left(1 - \frac{\eta}{2}\right)p \left(\frac{\eta}{2} - 0 \right) + (1 - p) \left(\frac{\eta}{2} - \frac{\frac{\eta}{2}(1-p)v}{(1-p)v + \left(1 - \frac{\eta}{2}\right)p(1-v)} \right) - \left(1 - \frac{\eta}{2}\right) \frac{\eta}{2}p \right) \\ &= v(1-p) \frac{\frac{\eta}{2} \left(1 - \frac{\eta}{2}\right)p(1-v)}{(1-p)v + \left(1 - \frac{\eta}{2}\right)p(1-v)} \end{aligned}$$

and

$$\begin{aligned} A_D^0 &= v \left(\Pr(d = \emptyset, X = -1 | q = 1, d = 0) \left(\Pr(\theta_h | \emptyset, X = -1, d = 0) - \frac{\eta}{2} \right) \right. \\ &\quad \left. + \Pr(d = \emptyset, X = 0 | q = 1, d = 0) \left(\Pr(\theta_h | \emptyset, X = 0, d = 0) - \frac{\eta}{2} \right) \right. \\ &\quad \left. - \frac{\eta}{2} \Pr(d = \theta_h) - \frac{\eta}{2} \Pr(d = \mu) \right) \\ &= v \left(\left(1 - \frac{\eta}{2}\right)p \left(\frac{\eta}{2 - \eta} - \frac{\eta}{2} \right) + (1 - p) \left(\frac{\frac{\eta}{2}((1-p)v + p(1-v))}{(1-p)v + \left(1 - \frac{\eta}{2}\right)p(1-v)} - \frac{\eta}{2} \right) - \frac{\eta}{2} \frac{\eta}{2}p \right) \\ &= v(1-p) \frac{\frac{\eta}{2}p(1-v)}{(1-p)v + \left(1 - \frac{\eta}{2}\right)p(1-v)}. \end{aligned}$$

Market Maker's Beliefs if the Manager Implements the Expansive Disclosure Policy

Following the same logic as before, we can calculate:

$$\Pr(\theta_h|\emptyset, X=0, d=1) = \frac{\Pr(\theta_h, d=\emptyset, X=0)}{\Pr(d=\emptyset, X=0)} = \frac{\frac{\eta}{2}(1-p)v}{(1-p)v + \frac{\eta}{2}p(1-v)} < \frac{\eta}{2}$$

$$\Pr(\theta_l|\emptyset, X=0, d=1) = \frac{\Pr(\theta_l, d=\emptyset, X=0)}{\Pr(d=\emptyset, X=0)} = \frac{\frac{\eta}{2}((1-p)v + p(1-v))}{(1-p)v + \frac{\eta}{2}p(1-v)} > \frac{\eta}{2}$$

If $X = -1$ and $d = \emptyset$, then from $X = -1$, the market maker can infer that the trader did not purchase the security because $s_{withheld} = 1$ and $s_{knowcash} = \mu$. Hence:

$$\Pr(\theta_h|\emptyset, X=-1, d=1) = 0$$

$$\Pr(\theta_l|\emptyset, X=-1, d=1) = 1.$$

Based on these beliefs and:

$$\Pr(d=\emptyset, X=-1|q=1, d=1) = \frac{\eta}{2}p$$

$$\Pr(d=\emptyset, X=0|q=1, d=1) = (1-p)$$

$$\Pr(d=\theta_h) = \frac{\eta}{2}p$$

$$\Pr(d=\mu) = (1-\eta)p$$

we can compute:

$$A_U^1 = u(1-p)\frac{\eta}{2} \frac{\frac{\eta}{2}p(1-v)}{(1-p)v + \frac{\eta}{2}p(1-v)}$$

and:

$$A_D^1 = u(1-p)\frac{\eta}{2} \frac{\left(1 - \frac{\eta}{2}\right)p(1-v)}{(1-p)v + \frac{\eta}{2}p(1-v)}.$$



Proof of Proposition 3**The Optimal Security Design $d = 0$ (OSD0) Program**

$$\min_{\Delta_D, \Delta_U} A_D^0 \Delta_D + A_U^0 \Delta_U \quad (13)$$

subject to:

$$pe \leq \Delta_D - (1-p)\Delta_U; \quad (14)$$

$$1 + e - \mu \leq \left(1 - \frac{\eta}{2} - A_D^0\right)\Delta_D + \left(\frac{\eta}{2} - A_U^0\right)\Delta_U; \quad (15)$$

$$0 \leq \Delta_i \leq e; \quad i \in \{D, U\}. \quad (16)$$

The feasible set of Δ_D and Δ_U is then determined by Equation (15) to be that subset of the square $[0, e] \times [0, e]$ in (Δ_D, Δ_U) -space lying below the line:

$$\Delta_U = \frac{\Delta_D - pe}{1-p} \quad (17)$$

and above the investor's "supply-financing" line:

$$\Delta_U = \frac{1 + e - \mu - \left(1 - \frac{\eta}{2} - A_D^0\right)\Delta_D}{\frac{\eta}{2} - A_U^0}. \quad (18)$$

We start by observing that, in all non-first best cases (i.e., when $1 + e - \mu > 0$), the investor's "supply financing" line (18) intersects the positive orthant in (Δ_D, Δ_U) -space. The key to determining the smallest value attained by the objective function (13) in the constraint set is the observation that the (downward-sloping) iso-cost lines of the objective function (13) are flatter than the slope of the line (18), as the following lemma states.

Lemma 3

$$\frac{A_D^0}{A_U^0} < \frac{1 - \eta/2 - A_D^0}{\eta/2 - A_U^0}.$$

Proof

Rewrite

$$A_D^0 \text{ and } A_U^0 \text{ as } A_D^0 = \frac{\eta}{2} A^0 \text{ and } A_U^0 = \left(1 - \frac{\eta}{2}\right) A^0,$$

where:

$$A^0 = v(1-p) \frac{\eta}{2} \frac{p(1-v)}{(1-p)v + \left(1 - \frac{\eta}{2}\right)p(1-v)}.$$

With this notation:

$$\frac{A_D^0}{A_U^0} = \frac{\frac{\eta}{2}A^0}{\left(1 - \frac{\eta}{2}\right)A^0} = \frac{\frac{\eta}{2}}{1 - \frac{\eta}{2}} \quad (19)$$

and $\frac{1-\eta/2-A\beta}{\eta/2-A\beta} = \frac{1-\frac{\eta}{2}-\frac{\eta}{2}A^0}{\frac{\eta}{2}-(1-\frac{\eta}{2})A^0}$. Then it is easy to confirm that the inequality in the statement of the lemma is equivalent to $0 < 1 - \eta$, which always holds. ■

Given this fact, it follows that the optimal security is defined by the “most southwest” point of the constraint set (14)–(16), as illustrated in Figure 2. We identify what that “most southwest” point is by studying the following four mutually exclusive and exhaustive cases (for $e > \mu - 1$).

- In case 1 (“low volatility”), the point on the line (18) lying on the Δ_D -axis, i.e., the point $(\Delta_D^0, \Delta_U^0) = \left(\frac{1+e-\mu}{1-\frac{\eta}{2}-A\beta}, 0\right)$, lies to the left of $(\Delta_D, \Delta_U) = (pe, 0)$.²¹ This case occurs when $\left(\frac{1+e-\mu}{1-\frac{\eta}{2}-A\beta} \leq pe\right)$ i.e., when $e \leq \frac{\mu-1}{1-(1-\frac{\eta}{2}-A\beta)p} \equiv e_{ID}^0$. In this case, the “southwest-most” (Δ_D, Δ_U) in the feasible set is $(pe, 0)$. This corresponds to the security $(\psi_I, \Delta_D, \Delta_U) = (\mu - e, pe, 0)$. This security is investment-grade debt with face value $\mu - (1 - p)e$.
- In case 2 (“medium volatility”), (Δ_D^0, Δ_U^0) lies to the right of $(\Delta_D, \Delta_U) = (pe, 0)$ but to the left of $(\Delta_D, \Delta_U) = (e, 0)$. This case occurs when $e > e_{ID}^0$ and $\frac{1+e-\mu}{1-\frac{\eta}{2}-A\beta} \leq e$, i.e., when $e_{ID}^0 < e \leq \frac{\mu-1}{\frac{\eta}{2}+A\beta} \equiv e_{SD}^0$. In this case, the “southwest-most” (Δ_D, Δ_U) in the feasible set is (Δ_D^0, Δ_U^0) . This corresponds to the security $(\psi_I, \Delta_D, \Delta_U) = \left(\mu - e, \frac{1+e-\mu}{1-\frac{\eta}{2}-A\beta}, 0\right)$. This security is investment-grade debt with face value $\mu - e + \frac{1+e-\mu}{1-\frac{\eta}{2}-A\beta}$.
- In case 3 (“high volatility”), (Δ_D^0, Δ_U^0) lies to the right of $(\Delta_D, \Delta_U) = (e, 0)$ and the point (Δ_D, Δ_U) on the line (18) evaluated at $\Delta_D = e$ lies inside the box $[0, e] \times [0, e]$. This case occurs when $e > e_{SD}^0$ and $\frac{1+e-\mu-(1-\frac{\eta}{2}-A\beta)e}{(\frac{\eta}{2}-A\beta)} \leq e$, so $e_{SD}^0 < e < \frac{\mu-1}{\frac{\eta}{2}+A\beta} \equiv \bar{e}_{SD}^0$. In this case, the “southwest-most” (Δ_D, Δ_U) in the feasible set is $\left(e, \frac{1+e-\mu-(1-\frac{\eta}{2}-A\beta)e}{\frac{\eta}{2}-A\beta}\right)$. This corresponds to the security $(\psi_I, \Delta_D, \Delta_U) = \left(\mu - e, e, \frac{1+e-\mu-(1-\frac{\eta}{2}-A\beta)e}{\frac{\eta}{2}-A\beta}\right)$. This security is speculative debt with face value $\frac{\mu(\frac{\eta}{2}-A\beta-1)+1+(\frac{\eta}{2}+A\beta)e}{\frac{\eta}{2}-A\beta}$.
- In case 4 (“extremely high volatility”), (Δ_D^0, Δ_U^0) lies to the right of $(\Delta_D, \Delta_U) = (e, 0)$, but the point (Δ_D, Δ_U) on the line (18) evaluated at $\Delta_D = e$ lies outside the box $[0, e] \times [0, e]$. This case occurs when $e \geq \bar{e}_{SD}^0$. In this case, no feasible security can implement the limited disclosure policy.

Substituting the values of the downside risk and upside potential of the securities that have been identified to be optimal in each of these four cases into the expression for the firm’s cost of capital in Proposition 1 completes the proof of Proposition 3. ■

²¹ Also note that $(\Delta_\mu^0, \Delta_\mu^0)$ lies to the right of $(\Delta_\mu, \Delta_\mu) = (0, 0)$ whenever first-best is not attainable.

Proof of Proposition 4

This proof is similar in style to Proposition 3, so its proof is omitted. ■

Proof of Proposition 5

We start the proof of the proposition by proving the following Lemma.

Lemma 4: For any $e > e_{ID}^1$ for which there exists a feasible security—call it “security 1”—that implements the expansive disclosure policy, there exists a feasible debt security that implements the limited disclosure policy that increases the manager’s expected profits relative to that obtained from security 1.

Proof: Suppose $e > e_{ID}^1$ and some security, say “security 1,” eliciting expansive disclosure exists. Then, we are in case 3 of Proposition 4. As the proof of Proposition 4 shows, optimal securities in this case are hybrid securities determined by those securities on the “supply-financing” line:

$$\Delta U = \frac{1 + e - \mu - \left(1 - \frac{\eta}{2} - A_D^1\right)\Delta_D}{\frac{\eta}{2} - A_U^1}$$

that lie in the constraint set for Program OSD1. Single out the optimal security on this line that also lies on the boundary line:

$$pe = \Delta_D - (1 - p)\Delta_U \quad (20)$$

of the constraint set. Call this latter security “security 2,” and denote its components by $(\psi_l^2, \Delta_D^2, \Delta_U^2)$. Security 2 necessarily generates at least as high expected profits for the manager as does security 1, because security 2 is optimal for Program OSD1 and security 1 is merely feasible for this program. Recall by Lemma 1 that securities that lie on the line (20) can induce the manager to adopt either the limited or expansive disclosure policy. In particular, this is true for security 2.

We now make two claims: security 2 is feasible for Program OSD0, and security 2 yields the manager higher expected profits under that program than under Program OSD1. This will prove that the manager is better off implementing the limited disclosure policy upon being given security 2 rather than the expansive disclosure policy. Notice that the firm’s cost of capital when the manager implements the expansive disclosure policy with security 2 satisfies:

$$\begin{aligned} A_D^1\Delta_D + A_U^1\Delta_U &= A_D^1(pe + (1 - p)\Delta_U) + A_U^1\Delta_U \\ &= (A_D^1 + A_U^1)\Delta_U + A_D^1p(e - \Delta_U) > (A_D^0 + A_U^0)\Delta_U + A_D^0p(e - \Delta_U) \\ &= A_D^0(pe + (1 - p)\Delta_U) + A_U^0\Delta_U = A_D^0\Delta_D + A_U^0\Delta_U, \end{aligned} \quad (21)$$

where the inequality follows from Corollary 1 and the first and last equalities follow from security 2 lying on (20). Of course (21) is equivalent to:

$$\left(1 - \frac{\eta}{2} - A_D^1\right)\Delta_D + \left(\frac{\eta}{2} - A_U^1\right)\Delta_U \leq \left(1 - \frac{\eta}{2} - A_D^0\right)\Delta_D + \left(\frac{\eta}{2} - A_U^0\right)\Delta_U. \quad (22)$$

(22), along with the feasibility of $(\psi_l^2, \Delta_D^2, \Delta_U^2)$ for Program OSD1 and the fact that $(\psi_l^2, \Delta_D^2, \Delta_U^2)$ lies on (20), implies the feasibility of $(\psi_l^2, \Delta_D^2, \Delta_U^2)$ for Program OSD0. (21) implies that the manager’s expected profits are strictly higher by implementing the limited disclosure policy than the expansive disclosure policy upon being given security 2. This proves the two claims.

To conclude the proof, recall that Proposition 3 established that debt securities always optimally implement the limited disclosure policy, so there is some debt security, say “security 3,” that is at least weakly preferable to security 2 that implements the limited disclosure policy. Since security 2 is weakly preferable to security 1, the proof of the lemma is complete. ■

Proof of Proposition 5

The proof of this proposition consists of the following three observations. First, the function defining the manager’s maximum expected profits when he implements the limited disclosure policy is linear in e over the interval $e \in (\mu - 1, e_{ID}^1]$; the same is also true of the function defining the manager’s maximum expected profits when he implements the expansive disclosure policy (for this same interval $(\mu - 1, e_{ID}^1]$).²² Second, the manager’s maximum expected profits are strictly higher under the expansive disclosure policy than under the limited disclosure policy at, and—by continuity—in the vicinity of, $e = \mu - 1$. (This follows from the discussion surrounding Proposition 2.) Third, the manager’s maximum expected profits are strictly higher under the limited disclosure policy than under the expansive disclosure policy at, and—by continuity—in the vicinity of $e = e_{ID}^1$. (This follows directly from Lemma 4 above.) Hence, these expected profit functions must cross exactly once in the interval $(\mu - 1, e_{ID}^1]$, which defines the point e_S in the statement of the proposition. The conclusion about the global optimality of risk-free debt, investment-grade debt, or speculative debt then follows directly from Propositions 3 and 4.

Proof of Proposition 7

Disclosure of θ_m is preferred over non-disclosure of θ_m when the following condition is satisfied:

$$\theta_m - \psi_m \geq E[\tilde{\theta} - \tilde{\psi} | d = \emptyset, d(\theta_m) = \emptyset],$$

where RHS of the inequality is given by:

$$E[\tilde{\theta} - \tilde{\psi} | d = \emptyset, d(\theta_m) = \emptyset] = \frac{p(1 - \eta)(\theta_m - \psi_m) + (1 - p)\left(\theta_m - \left(\frac{\eta}{2}\psi_h + (1 - \eta)\psi_m + \frac{\eta}{2}\psi_l\right)\right)}{p(1 - \eta) + (1 - p)}$$

Simplification yields:

$$\theta_m - \psi_m \geq E[\tilde{\theta} - \tilde{\psi} | d = \emptyset, d(\theta_m) = \emptyset]$$

$$\theta_m - \psi_m \geq \theta_m - \frac{\psi_h + \psi_l}{2}$$

$$\frac{\psi_h + \psi_l}{2} \geq \psi_m$$

²² Because: (1) by Proposition 3, the manager’s expected profit function when he adopts the limited disclosure policy is linear in e for e in the interval $(\mu - 1, e_{ID}^0]$; (2) by Proposition 4, the manager’s expected profit function is linear in e for e in the interval $(\mu - 1, \bar{e}_{ID}^1]$, and so, since $e_{ID}^1 < \bar{e}_{ID}^1$, *a fortiori* is linear over the interval $(\mu - 1, e_{ID}^1]$; (3) $e_{ID}^1 = \frac{\mu - 1}{1 - (1 - \frac{\eta}{2} - A_D^1)p} < \frac{\mu - 1}{1 - (1 - \frac{\eta}{2} - A_D^0)p} = e_{ID}^0$, because $A_D^0 < A_D^1$, by Corollary 1 above; and so (4) both of these expected profit functions are linear in e for e in the interval $(\mu - 1, e_{ID}^1]$.

$$\Delta_U \geq \Delta_D.$$

The case for which non-disclosure of θ_m is preferred over disclosure of θ_m is identical.

Next, we derive the trader's optimal trading strategy as a function of the manager's disclosure policy. Suppose the manager discloses both the low and high state but not the intermediate state. Then, the trader prefers to buy the security when $s_{withheld} = 0$ if and only if her expected purchasing price, $uP(\emptyset, 0) + (1 - v)P(\emptyset, 1)$, is less than the expected payoff of the security, $E[\tilde{\psi}]$. $P(\emptyset, 1)$ equals $E[\tilde{\psi}]$ (because the market maker can infer that $s_{withheld} = 0$) and $P(\emptyset, 0)$ is a weighted average of $E[\tilde{\psi}]$ and ψ_m . Hence, $y(\emptyset, 0) = 1$ is optimal if $E[\tilde{\psi}] > \psi_m$ or $\Delta_U > \Delta_D$. This condition is incompatible with the manager choosing not to disclose θ_m . Hence, the trader does not buy the security when $s_{withheld} = 0$. The trader prefers to buy the security when $s_{withheld} = 1$ if and only if her expected purchasing price, $uP(\emptyset, 0) + (1 - v)P(\emptyset, 1)$, is less than the payoff of the security, ψ_m . $P(\emptyset, 1)$ equals ψ_m (because the market maker can infer that $s_{withheld} = 1$) and $P(\emptyset, 0)$ is a weighted average of $E[\tilde{\psi}]$ and ψ_m . Hence, $y(\emptyset, 1) = 1$ is optimal if $\psi_m > E[\tilde{\psi}]$ or $\Delta_U < \Delta_D$. This condition is consistent with the manager choosing not to disclose θ_m . Hence, the trader buys the security when $s_{withheld} = 1$. Clearly, the trader never buys the security when the manager fully discloses his private information. ■

REFERENCES

- Bias, B., T. Mariotti, G. Plantin, and J.-C. Rochet. 2007. Dynamic security design: Convergence to continuous time and asset pricing implications. *Review of Economic Studies* 74 (2): 345–390.
- Botosan, C. A. 1997. Disclosure level and the cost of equity capital. *The Accounting Review* 72 (3): 323–349.
- , and M. A. Plumlee. 2002. Re-examination of disclosure level and the expected cost of equity capital. *Journal of Accounting Research* 40 (1): 21–40.
- Casamatta, C. 2003. Financing and Advising: Optimal financial contracts with venture capitalists. *Journal of Finance* 58 (5): 2059–2086.
- Christensen, P., G. Feltham, and M. Wu. 2002. Cost of capital in residual income for performance evaluation. *The Accounting Review* 77 (1): 1–23.
- Demange, G., and G. Laroque. 1995. Private information and the design of securities. *Journal of Economic Theory* 65 (1): 233–257.
- DeMarzo, P., and D. Duffie. 1999. A liquidity-based model of security design. *Econometrica* 67 (1): 65–99.
- Dye, R. A. 1985. Disclosure of nonproprietary information. *Journal of Accounting Research* 23 (1): 123–145.
- . 1998. Investor sophistication and voluntary disclosures. *Review of Accounting Studies* 3 (3): 261–287.
- Fulghieri, P., and D. Lukin. 2001. Information production, dilution costs, and optimal security design. *Journal of Financial Economics* 61 (1): 3–42.
- Harris, M., and A. Raviv. 1990. Capital structure and the informational role of debt. *The Journal of Finance* 45 (2): 321–349.
- , and ———. 1991. The theory of capital structure. *The Journal of Finance* 46 (1): 297–356.
- Innes, R. 1990. Limited liability and incentive contracting with ex-ante action choices. *Journal of Economic Theory* 52: 45–67.
- Jorgensen, B., and M. Kirschenheiter. 2003. Discretionary risk disclosures. *The Accounting Review* 78 (2): 449–469.
- Jung, W.-O., and Y. K. Kwon. 1988. Disclosure when the market is unsure of information endowment of managers. *Journal of Accounting Research* 26 (1): 146–153.
- Kanodia, C. S., A. Mukherji, H. Sapra, and R. Venugopalan. 2000. Hedge disclosures, futures prices and production distortions. *Journal of Accounting Research* 38 (Supplement): 53–82.
- Korajczyk, R. A., D. J. Lucas, and R. L. McDonald. 1991. The effect of information releases on the pricing and timing of equity issues. *Review of Financial Studies* 4 (4): 685–708.
- , ———, and ———. 1992. Equity issues with time-varying asymmetric information. *Journal of Finan-*

- cial and Quantitative Analysis* 27 (3): 397–417.
- Kyle, A. 1985. Continuous auctions and insider trading. *Econometrica* 53 (6): 1315–1335.
- Myers, S., and N. S. Majluf. 1984. Corporate financing and investment decisions when firms have information that investors do not have. *Journal of Financial Economics* 13 (2): 187–221.
- Nachman, D. C., and T. H. Noe. 1994. Optimal design of securities under asymmetric information. *Review of Financial Studies* 7 (1): 1–44.
- Pae, S. 2002. Discretionary disclosure, efficiency, and signal informativeness. *Journal of Accounting and Economics* 33 (3): 279–311.
- Rahi, R. 1995. Partially revealing rational expectations equilibria with nominal assets. *Journal of Mathematical Economics* 24 (2): 137–146.
- Repullo, R., and J. Suarez. 2004. Venture capital finance: A security design approach. *Review of Finance* 8 (1): 75–108.
- Stocken, P. C. 2000. Credibility of voluntary disclosure. *The RAND Journal of Economics* 31 (2): 359–374.
- Sunder, J. 2006. Information spillovers and capital structure. Working paper, Northwestern University.

Dividends, Share Repurchases, and Tax Clienteles: Evidence from the 2003 Reductions in Shareholder Taxes

Jennifer L. Blouin

University of Pennsylvania

Jana S. Raedy

The University of North Carolina at Chapel Hill

Douglas A. Shackelford

*The University of North Carolina at Chapel Hill and
National Bureau of Economic Research*

ABSTRACT: This study jointly evaluates firm-level changes in investor composition and shareholder distributions following a 2003 reduction in the dividend and capital gains tax rates for individuals. We find that directors and officers, but not other individual investors, rebalanced their portfolios to maximize after-tax returns in light of the new tax rules. We also find that firms adjusted their distribution policy (specifically, dividends versus share repurchases) in a manner consistent with the altered tax incentives for individual investors. To our knowledge, this is the first study to employ simultaneous equations to estimate both shareholder and managerial responses to the 2003 rate reductions. We find that the generalized method of moments (GMM) estimates are substantially stronger than OLS estimates, consistent with our expectation that investor and manager responses are simultaneously determined. Failure to estimate systems of equations may account for some of the weak and conflicting results from prior studies of the 2003 rate reductions.

Keywords: *shareholder taxes; dividends; share repurchases; tax clienteles.*

Data Availability: *Data used in this study are available from public sources identified in the study.*

The authors have benefited greatly from comments from Dan Dhaliwal and two anonymous referees, discussions with Jeff Brown, Raj Chetty, Emmanuel Saez, and Scott Weisbenner, and research assistance from Kevin Markle. We are responsible for all remaining errors.

Editor's note: Accepted by Steven Kachelmeier, with thanks to Dan Dhaliwal for serving as editor on a previous version.

*Submitted: November 2007
Accepted: September 2010
Published Online: April 2011*

I. INTRODUCTION

This study employs simultaneous equations to assess the responses of both investors and companies to the 2003 reductions in dividend and capital gains tax rates. We predict that individual investors, the only ones affected by the reduction in shareholder taxes, rebalanced their portfolios to maximize after-tax returns in light of the new tax rules. We also predict that firms adjusted their distribution policies (specifically, dividends versus share repurchases) to maximize share value—i.e., distributing profits in a manner that was most attractive to their investors after considering shareholder taxes. With regard to investor responses, we find evidence that insiders (i.e., directors and officers) increased holdings in their own companies if their dividend-repurchase mix reflected the new tax incentives. However, we find no evidence that other individual investors rebalanced their portfolios. With regard to managerial responses, we find that firms with disproportionately large individual holdings modified their payouts in a manner consistent with the altered tax incentives. However, changes in dividend and repurchase policy were not immediate; firms deferred widespread, substantial changes until the second quarter following enactment.

To our knowledge, this is the first study to jointly estimate investor and firm responses to changes in shareholder taxes. The fact that both investors and firms can change their behavior following a change in shareholder taxes presents an identification problem. To illustrate, suppose we test for an association between dividend yields and individual stock ownership and find that the correlation becomes more positive following a reduction in individual dividend tax rates. Such a finding is consistent with both (1) individuals switching to high-dividend-paying firms following the tax cut (i.e., a tax clientele response) and (2) firms that are held mostly by individuals increasing their dividends following the tax reduction (a firm payout response). To distinguish between investors rebalancing their portfolios and firms altering their distributions, this study estimates simultaneous equations. Prior studies have focused on either investors or managers, but not both, and results have been weak and in conflict. By adopting simultaneous equations, we find evidence of substantially stronger investor and managerial responses to a decline in shareholder tax rates.

To maximize the power of our tests, we compare a firm's ownership and dividend-repurchase mix before and after the largest change in U.S. dividend taxation. The Jobs and Growth Tax Relief Reconciliation Act of 2003 (JGTRRA) reduced, for individuals only, the maximum, statutory dividend tax rate from 38.6 percent to 15 percent. It also lowered the maximum, statutory individual capital gains tax rate, which applies to share repurchases, from 20 percent to 15 percent. In many ways, JGTRRA is an ideal legislative setting for testing an association between shareholder taxes and payout policy. The scope of JGTRRA was narrow. Its genesis was individual dividend tax reduction. Its primary amendment was individual capital gains tax relief.¹ The final bill did little more than reduce dividend and capital gains tax rates. Nonetheless, the economic effects were huge.² Because of its narrow focus and big impact, the JGTRRA provides a much stronger setting than other shareholder tax rate changes, which were either much smaller (e.g., Tax Relief Act of 1997) or involved widespread overhaul of the tax system that affected far more than just shareholder taxes (e.g., Tax Reform Act of 1986).

The JGTRRA dividend and capital gains tax cuts should have altered the optimal mix of dividends and repurchases for at least some individual investors. As a result, we expect that some individuals rebalanced their portfolios so that a higher proportion of their equity returns came in

¹ For the legislative history, see Auerbach and Hassett (2007).

² The size of the tax savings has been very large. In September 2007, the U.S. Congress' Joint Committee on Taxation reported that the 2003 dividend and capital gains tax rates will cost the Treasury \$632 billion between 2007 and 2011, the largest tax expenditure in the tax code.

the form of dividends. Consistent with such portfolio rebalancing, Desai and Dharmapala (2011) report that total U.S. equity investments shifted from foreign countries, whose companies did not qualify for the lower dividend tax rates, to foreign countries, whose companies did qualify for the lower dividend tax rates. However, to our knowledge, no one has documented whether individual investors rebalanced their much larger holdings of domestic holdings in a manner consistent with the changed tax incentives following passage of the JGTRRA. This study addresses that void by studying the domestic portfolio rebalancing of three different individual investor groups.

To attract individual investors who were looking for more dividends following passage of the legislation, we expect that at least some firms increased the dividend portion of their distributions. Consistent with such a managerial response, Chetty and Saez (2005) document that dividend initiations jumped in 2003. Brown et al. (2007) find that reductions in share repurchases funded these 2003 dividend initiations, with such substitution limited to companies where directors and officers held disproportionately large shares. They find no similar managerial responses among the set of firms that account for almost all dividend issuances—i.e., the firms paying dividends before JGTRRA, or among firms without high insider ownership.³ Comparing executive compensation in 2003 and 2002, Aboody and Kasznik (2008) also reach mixed conclusions about the changes in the dividend-repurchase mix.

We conduct a more comprehensive study of managerial responses to JGTRRA by looking at dividend-paying firms (which far exceed initiators in number and payout), studying both insiders and other individual investors, and extending the analysis to include changes in distribution policies through 2005 (two years beyond the extant JGTRRA studies). Lengthening the investigation period enables us to calibrate how long it took investors and managers to respond to JGTRRA. We also estimate a system of equations and find that results are statistically and economically more significant under simultaneous equations than under ordinary least-squares.

To test for investor and firm responses, we compare the percentage of shares held before and after JGTRRA for three groups of individual investors: insiders (i.e., directors and officers of the firm), other non-executive individuals investing on their own account, and mutual fund investors.⁴ We aggregate each firm's dividends and repurchases during the eight quarters immediately preceding the quarter of enactment and compute the ratio of dividends to total payout (dividends plus repurchases).⁵ We compare that ratio to the one formed with aggregated dividends and repurchases for the eight quarters immediately after the quarter of enactment. We then test for an association between the change in shareholders and change in the dividend-repurchase mix. This difference-in-differences approach mitigates the likelihood of spurious conclusions arising from omitted correlated variables. Estimating a system of equations enables us to determine whether the association is driven by tax clientele effects, payout changes, or both.

We find evidence consistent with both investor and firm responses to JGTRRA. However, insiders are the only investors who appear to have altered their holdings in response to JGTRRA. We find stronger evidence that firms modified their payout policy in response to changes in

³ Brown et al. (2007) are careful to state that their inferences about substitution are limited to 2003 dividend initiators. That said, because their sample includes both initiators and non-initiating firms, it was unclear to us whether the findings reported in Table 6 of their study shed any light on the firm responses of non-initiators. Private conversations with the authors confirmed that their conclusions were limited to 2003 dividend initiators. In addition, to enable us to investigate the non-initiators in their study more closely, the authors kindly provided us with their data, for which we were most appreciative. We replicated their results, and, consistent with the inferences in their study and our conversations with them, we found that their results only hold for firms that initiated in 2003.

⁴ Many mutual fund investors are not individuals and the earnings for many mutual funds held by individuals are not subject to JGTRRA, e.g., 401(k) investments. Unfortunately, we cannot observe the extent to which JGTRRA affects a specific mutual funds' investor base. This limitation biases against our finding a response by mutual fund investors.

⁵ We aggregate to reduce some of the noise arising because repurchases, unlike regular quarterly dividends, are uneven and irregular.

shareholder taxes. In particular, we find that the movement toward distributing a larger proportion of profits as dividends was greatest among not only those companies held disproportionately by individual investors, particularly directors and officers, but also those firms where other individuals and mutual funds had large holdings. We find that firms began to substantially alter their distribution policies in the second quarter following the passage of JGTRRA, consistent with firms needing time to adjust their dividend and repurchase policies.

We find that the economic significance of the 2003 legislation is substantially greater using simultaneous equations than using ordinary least-squares (OLS). This is an important contribution of this study because, to our knowledge, we are the first study to jointly evaluate shareholder and firm behavior. Based solely on our OLS results, we would erroneously conclude that no investors rebalanced their portfolios and that the firm responses were limited to those companies with heavy insider ownership. Using generalized method of moments (GMM) estimation of the system of equations, which is the appropriate approach given our expectation that investors and managers simultaneously readjusted their behavior after 2003, we find that some investors, namely individuals, altered their portfolios and that firms were responsive to insiders, mutual funds, and other individual tax considerations.

The next section develops testable hypotheses. Section III details the research design. Section IV presents the empirical findings. Closing remarks follow in Section V.

II. HYPOTHESIS DEVELOPMENT

Framework

This study does not attempt to tackle the longstanding puzzle of why firms pay dividends when they could distribute profits through share repurchases, which remain tax-advantaged, though less so, even after JGTRRA.⁶ Rather, in a nutshell, we: (1) accept the fact that investors desired and some firms paid dividends before JGTRRA (obviously for non-tax reasons), (2) assume that the mix of dividends and repurchases was optimal before passage of the legislation, (3) expect that the large tax rate reductions for dividends (compared with the relatively modest reductions for capital gains) led some individuals to rebalance their portfolios in favor of dividend income, and (4) predict that some firms, in response to the changing tax incentives for individual investors, increased the portion of their distributions in the form of dividends after JGTRRA. The remainder of this section elaborates on these relations to develop formal hypotheses about the impact of JGTRRA on shareholder distributions.

To develop the intuition for our hypotheses, we start with a simple framework.⁷ Absent taxes, suppose that all investors hold optimally diversified portfolios of the risk-free asset and the market portfolio. In that setting, shareholder ownership of stocks should not vary across investors with similar risk preferences. Now suppose some investors (call them individuals) become taxed on their dividend income. These individual investors will no longer hold the original, pre-tax, optimally diversified portfolio. Rather, they will underweight their portfolio in tax-disfavored, dividend-paying stocks and overweight their portfolio in tax-favored, no-dividend stocks. This shift will boost the price of no-dividend stocks and drive down the price of dividend-paying

⁶ For a sampling of the "dividend puzzle" literature, see Miller and Modigliani (1961), Feldstein and Green (1983), Bagwell and Shoven (1989), Bertheim (1991), DeAngelo et al. (2006), Jagannathan et al. (2000), Fama and French (2001), Grullon and Michaely (2002), Dhaliwal and Li (2006), Gordon and Dietz (2008), and Chetty and Saez (2010), among many others. On a different note, even though both repurchases and dividends now face a maximum tax rate of 15 percent, repurchases remain tax-advantaged for at least two reasons. First, sellers can offset the tax basis of the shares that they sell against the proceeds from the sale in computing their capital gains. Second, they can offset those capital gains with capital losses that otherwise might not be deductible. This contrasts with dividends, where the entire amount is taxed upon receipt.

⁷ We appreciate the contributions of an anonymous referee in sketching out this framework.

stocks. These price movements will entice non-individual investors (call them tax-exempts) to hold more dividend stocks and less no-dividend stocks. As a result, each investor ultimately would hold the portfolio that features his optimal trade-off between risk and after-tax return; the share price of each stock would equate supply and demand; and a heterogeneous mix of shareholders would emerge endogenously. One example of an empirical study of such tax clientele responses is Dhaliwal et al. (1999), who document increases in institutional holdings when firms initiate dividends, which are tax-disadvantaged, to individual investors.⁸

Similarly, if individuals were taxed more heavily on dividends than on share repurchases, then they would overweight their portfolios with stocks that distributed disproportionate amounts of profits through repurchases, as compared to dividends.⁹ Meanwhile, non-individual investors would overweight their portfolios with stocks that distributed disproportionate amounts of profits through dividends, as compared to share repurchases. In this setting, managers would choose both the level and mix (dividends versus share repurchases) of shareholder distributions that maximize the firm's stock price. Their distribution policy would affect both the equilibrium price and the equilibrium mix of investors and changes in their payout policy will induce a change in the mix of investors that own the firm.¹⁰

Now consider a reduction in the dividend tax rate. Individual investors would rebalance their portfolios, adding more dividend-paying stocks than was optimal under the prior high-dividend tax regime. This will drive up the price of dividend-paying stocks, altering the optimal portfolio mix for tax-exempt investors who would now shift from dividend-paying stocks to no-dividend stocks at the margin. At the same time that investors would be adjusting their portfolios in light of the changes in dividend tax policy, managers would be revising their level and mix of shareholder distributions in light of the new tax policy to continue to maximize their stock price. As a result, both the investor mix (individuals versus tax-exempts) and the firm's distribution policy (dividends versus share repurchases) would change after a reduction in the dividend tax rate.

Determining the relative importance of the investor response and the firm's response would require joint evaluation of both investors' and managers' incentives. Studies that examine only individual investors' new tax-motivated demand for dividends might erroneously attribute all of the increase in dividends to a clientele effect. In the extreme, this is true even if investors did not rebalance their portfolio, but rather, firms simply increased their dividend payouts. Likewise, studies (such as several prior examinations of JGTRRA) that focused solely on firm's new tax-driven supply of dividends might erroneously attribute increased dividend income by individuals to a payout response when the result was actually due to individual investors rebalancing their portfolios. In the tests below, we find evidence of both a clientele and a managerial response to shareholder tax rate reductions.

Investor Responses

We begin our hypothesis development by focusing on potential investor responses to JGTRRA. We assume that because the decline in the dividend tax rate (from 38.6 percent to 15

⁸ Moser and Puckett (2009) include the JGTRRA in their study of whether there is a positive association between the portion of dividend-paying securities in tax-advantaged institutions' portfolios and the dividend penalty. Although they find evidence of a positive association, their analysis is confounded by the classification of financial institutions as tax-advantaged institutions (Blouin 2009).

⁹ We are assuming that tax-sensitive investors do not interpret any potential dividend and repurchase signals differently from other market participants.

¹⁰ For a formal equilibrium model of stock ownership and stock price, see Guenther and Sansing (2010). Motivated by recent empirical findings of a relation between tax-exempt ownership and dividend tax capitalization, they model and provide empirical evidence that dividend tax capitalization reflects all investors' weighted average tax rates, where the weighting varies with the investors' risk tolerances. An important difference between their study and ours is that they focus on prices and ownership, whereas we focus on payouts and ownership.

percent) exceeded the decrease in the capital gains tax rate (from 20 percent to 15 percent), the net effect of JGTRRA was to make firms that distributed profits mostly through dividends more attractive for individual investors than those that distribute profits mostly through share repurchases. Assuming investors were holding the optimal portfolio (considering risk and taxes) before JGTRRA's enactment, we predict that, after enactment, individual investors altered their portfolios to receive a higher percentage of their returns in the form of dividends.

H1: Individual investors responded to JGTRRA by increasing their holdings in stocks that distribute larger proportions of their profits through dividends.

H1 notwithstanding, it is important to recognize that investors cannot freely rebalance their portfolios. Aside from the commissions and other transaction costs that investors face on all trades, taxable investors pay capital gains taxes on any excess of the proceeds from the sale of the stock over the basis of the stock. Thus, some investors may accept an inferior portfolio, rather than incur the tax and non-tax costs of rebalancing their portfolio. This is one reason why we might not observe investors engaging in widespread rebalancing following JGTRRA.

Firm Responses

We anticipate that firms will respond to individual investors' enhanced interest in dividends by increasing the portion of profits that they distribute in the form of dividends. Those firms wishing to retain or increase their holdings by individual investors are likely to distribute more of their profits as dividends, following the enactment of JGTRRA, than other firms. Unfortunately, we cannot observe the pool of future shareholders that firms hope to attract by adjusting their payout policy. Therefore, we look to their shareholder mix at passage and assume that firms with greater individual ownership at enactment would be more likely to alter their distribution policy to distribute more of their profits as dividends than those firms with less individual ownership. To the extent a firm's current investor mix is not a good predictor of managers' desired investor mix in the future, our tests are biased against finding a firm response to JGTRRA.

H2: Managers responded to JGTRRA by distributing a larger portion of their profits as dividends. The extent to which managers increased their dividend percentage increased with the individual ownership of their firm.

The tension in H2 arises from several reasons why managers might not have adjusted their dividend-repurchase mix in response to JGTRRA. First, the JGTRRA tax rate reductions were scheduled to expire in five years at the end of 2008 (later deferred to the end of 2010), and 2004 Democratic Presidential candidate, John Kerry, pledged to restore the higher dividend tax rates for the two highest tax brackets, if elected. Since dividends tend to be sticky and the markets historically punish firms for decreasing dividends, many firms may have chosen to leave their distribution policy unchanged, delay any change until after the 2004 elections, or turn to one-time special dividends (which we ignore in this study). Second, dividends are purported to play an important role by alleviating asymmetric information through conveying private information to the market. Tax-motivated dividend changes might undermine this signal. Third, large increases in dividends could adversely affect the firm's compensation structure, particularly to the extent that the firm relies on stock options, which are not dividend-protected. Consistent with this deterrent to modifying distributions, Aboody and Kasznik (2008) find that firms that increased their dividends after JGTRRA modified their stock option and restricted stock compensation plans. In short, there were multiple reasons for firms to be hesitant about modifying their distribution policy following JGTRRA, even if individuals were seeking higher dividends than before the legislation.

III. RESEARCH DESIGN

System of Regression Equations

As discussed above, we jointly evaluate the changes in investor composition and distribution policy following the enactment of JGTRRA. We predict that individual investors rebalanced their portfolios so that a larger proportion of their shareholder income was in the form of dividends as opposed to share repurchases. We also expect that managers altered their distributions so that a larger proportion of their distributions were dividends and that this adjustment was increasing in the extent of individual ownership.

To jointly evaluate the impact of investor and manager responses to JGTRRA, we estimate a system of four equations (variables are defined below):

$$\begin{aligned} INSIDER = & \alpha_0 + \alpha_1 * POST + \alpha_2 * DIV\% + \alpha_3 * DIV\% * POST + \alpha_4 * S\&PRATING \\ & + \alpha_5 * AGE + \alpha_6 * SP500 + \alpha_7 * LIQUIDITY + \alpha_8 * BETA + \alpha_9 * IRISK \\ & + \alpha_{10} * MKTADJRET + \alpha_{11} * SALESGR + \alpha_{12} * R\&DINT; \end{aligned} \quad (1)$$

$$\begin{aligned} NONEXEC = & \delta_0 + \delta_1 * POST + \delta_2 * DIV\% + \delta_3 * DIV\% * POST + \delta_4 * S\&PRATING \\ & + \delta_5 * AGE + \delta_6 * SP500 + \delta_7 * LIQUIDITY + \delta_8 * BETA + \delta_9 * IRISK \\ & + \delta_{10} * MKTADJRET + \delta_{11} * SALESGR + \delta_{12} * R\&DINT; \end{aligned} \quad (2)$$

$$\begin{aligned} MF = & \gamma_0 + \gamma_1 * POST + \gamma_2 * DIV\% + \gamma_3 * DIV\% * POST + \gamma_4 * S\&PRATING + \gamma_5 * AGE \\ & + \gamma_6 * SP500 + \gamma_7 * LIQUIDITY + \gamma_8 * BETA + \gamma_9 * IRISK + \gamma_{10} * MKTADJRET \\ & + \gamma_{11} * SALESGR + \gamma_{12} * R\&DINT; \end{aligned} \quad (3)$$

$$\begin{aligned} DIV\% = & \beta_0 + \beta_1 * POST + \beta_2 * INSIDER + \beta_3 * INSIDER * POST + \beta_4 * NONEXEC \\ & + \beta_5 * NONEXEC * POST + \beta_6 * MF + \beta_7 * MF * POST + \beta_8 * RE + \beta_9 * PERM \\ & + \beta_{10} * TRANS + \beta_{11} * FCF + \beta_{12} * DYIELD + \beta_{13} * LEVERAGE + \beta_{14} * SIZE \\ & + \text{Industry Indicators.} \end{aligned} \quad (4)$$

To capture investor responses, the first three equations regress the percentage of the firm held by three individual investor groups on the percentage of the firm's payouts that are dividends (*DIV%*). There is one equation for each investor group: insiders (*INSIDER* in Equation (1)), other individuals (*NONEXEC* in Equation (2)), and mutual funds (*MF* in Equation (3)). The coefficients on the key variables of interest are α_3 , δ_3 , and γ_3 , respectively. The fourth equation in the system flips the direction of the association and regresses the percentage of the firm's payouts that are dividends on each of the three investor groups. The fourth equation has three key variables of interest, one for each investor group in the period post-enactment: insiders (β_3), other individuals (β_5), and mutual funds (β_7).

Dividends-to-Payout Ratio Variable

The percentage of the firm's payouts that are dividends (*DIV%*) is an explanatory variable in the first three equations and the dependent variable in the final equation. *DIV%* is a ratio where the numerator is the sum of dividends over an eight-quarter period and the denominator is the sum of dividends and share repurchases over the same eight quarters. There are two eight-quarter periods for each firm. The first eight quarters are the eight quarters immediately before the fiscal quarter

in which JGTRRA was enacted.¹¹ The second eight quarters are the eight quarters immediately following the fiscal quarter in which the JGTRRA was enacted.¹² Consequently, each firm has two *DIV%* measures: one before enactment and one after enactment.

We aggregate distributions over a two-year period because, unlike regular, quarterly dividends, share repurchases are irregular. Therefore, focusing on the dividend-repurchase mix in a single quarter could introduce excessive noise. That said, in sensitivity tests reported later, we relax this aggregation requirement and report results on a quarterly basis. Inferences are largely unaltered.

We can measure repurchases in two ways. One option is total share repurchases. Another option is net repurchases—i.e., total share repurchases less stock issuances.¹³ We use net repurchases because we are interested in the cash that the firm could have distributed as dividends. Fama and French (2001) note that dividends cannot substitute for repurchases in many situations. Firms need shares for executive compensation, stock option exercises, stock acquisitions, and funding employee stock ownership plans, among other things. Thus, consistent with Fama and French (2001), we measure net repurchases as the change in treasury stock.¹⁴ If there is a net decrease in treasury stock, then we truncate our measure of repurchases at zero. For those firms that do not use the treasury stock method, we measure net repurchases as total repurchases from the statement of cash flows less decreases in preferred stock.

For post-enactment observations in the first three regressions, we interact *DIV%* with *POST*, a categorical variable that equals 1 for observations after the May 23, 2003 enactment (i.e., quarters after the second quarter of 2003). In these three tax clientele tests, positive coefficients on *DIV% * POST* (the coefficients are α_3 , δ_3 , and γ_3 in the system of equations) can be interpreted as evidence that, after passage of the JGTRRA, individuals rebalanced their portfolios by shifting toward stocks where dividends constituted a larger portion of total payouts. We also include *POST* as a separate variable in each regression equation to capture any other temporal change.

Investor Group Variables

We employ three variables to capture the portion of the firm owned by individual investors. The first individual ownership measure, *INSIDER*, is the percentage of shares held by directors and officers, as reported in Thomson Financial's Insider Filing Data.¹⁵ Note that these shareholders play dual roles—as the managers setting distribution policy and as individual shareholders, often

¹¹ We exclude the enactment quarter (May 2003) because it is unclear which tax regime managers were contemplating when they issued dividends and repurchased shares during that quarter. We treat the first quarter of 2003 as a pre-enactment quarter, even though the legislation was retroactive to the beginning of the year. The reason is that passage of the legislation was uncertain until Vice-President Cheney's tiebreaking vote in the U.S. Senate in May. Sensitivity tests, detailed below, provide assurance that this classification is appropriate. In addition, inferences hold if we exclude any 2003 quarters from the pre-enactment period.

¹² Blouin and Krull (2009) show that share repurchases rose in 2005 as firms enjoyed a tax holiday for repatriating earnings from foreign subsidiaries. When we replicate our analysis excluding the 2005 quarters, inferences hold.

¹³ For further detail, see the discussion in footnotes 5 and 6 of Boudoukh et al. (2007) and footnote 9 of Skinner (2008).

¹⁴ Using treasury stock to measure repurchases is not without limitations. As Fama and French (2001) point out, using annual changes in treasury stock will fail to match a repurchase in one year and its reissuance in another year. This problem is mitigated in our research design because we combine two years of activity into one observation. However, even aggregation over two years cannot fully eliminate the potential mismeasurement.

¹⁵ The reporting of holdings of insiders is mandated by Section 16 of the Securities Exchange Act of 1934, which applies to every person who is the beneficial owner of more than 10 percent of any class of equity security registered under Section 12 of the Exchange Act and each director and officer (collectively, "reporting persons" or "insiders") of the issuer of the security. On a different note, conclusions do not change if we limit *INSIDER* to direct holdings, excluding shares held by family members, trusts and corporations controlled by the insider, and similar related parties.

with large stockholdings and suffering from inadequate diversification.¹⁶

The second measure of individual ownership is *NONEXEC*, which is intended to measure all individual holdings, other than those by insiders or through mutual funds. Ideally, we would measure the number of shares for which dividends and capital gains are expected to flow through to individual tax returns—i.e., those shares held by individuals or flow-through entities (e.g., mutual funds, partnerships, trusts, S corporations, or limited liability corporations) whose income is reported on U.S. individual tax returns. This ideal measure would exclude those holdings for which the dividends and capital gains do not flow through to individual tax returns, such as tax-exempt organizations, corporations, foreigners, and tax-deferred accounts (e.g., qualified retirement plans, including pensions, 401(k), and IRAs). Unfortunately, the ideal measure does not exist. Thus, as in Ayers et al. (2002), Ayers et al. (2004), Blouin et al. (2003), and Dhaliwal et al. (2003), among others, we use 13-F filings to estimate the percentage of the firm held by individual shareholders. *NONEXEC* is 1 less (1) the percentage of shares that institutional investors own, as reported in 13-F filings and collected by Thomson Financial's Institutional Holdings database, (2) the percentage of shares held by non-officer/director beneficial owners as reported in Thomson Financial's Insider Filing data, and (3) *INSIDER*.

The third measure of individual ownership, *MF*, is the percentage of the firm owned by mutual funds, as reported in 13-F filings and collected by Thomson Financial's Institutional Holdings database. As mentioned above, this is an imperfect measure of individual ownership because mutual funds include both investments that are subject to personal taxes and investments that are not subject to personal taxes. Sometimes the dividends and capital gains realized by mutual funds are taxed at the individual level. At other times, distributions to mutual funds are exempt because the shares are held in deferred tax accounts, such as 401(k) or IRAs. We include *MF* as an attempt to capture all shareholder income that is taxed on personal tax returns. However, we recognize that the unobservable measurement error in *MF* (arising from the inclusion of non-individual owners) may lessen its usefulness.

Each of the three individual ownership measures serves as the dependent variable in one of the first three regressions. All three are explanatory variables in the fourth regression, which tests for firm responses. Positive coefficients on these variables are consistent with firms altering their payouts in response to changes in individual tax incentives. Specifically, a positive coefficient on *INSIDER * POST* (β_3) in the fourth regression will be interpreted as evidence that, after passage of the JGTRRA, the percentage of payouts distributed as dividends was increasing in insider ownership. It seems likely that, when directors and officers hold large shares of a firm, payouts are likely to be particularly responsive to individual tax incentives. A positive coefficient on *NON-EXEC * POST* (β_5) in the fourth regression will be interpreted as evidence that, after passage, *DIV%* was increasing in the percentage of the firm held by non-executive shareholders. A positive coefficient on *MF * POST* (β_7) in the fourth regression will be interpreted as evidence that, after passage, *DIV%* was increasing in the percentage of the firm held by mutual funds. The measurement error in *MF* (arising from the fact that non-individuals invest in mutual funds) should bias the coefficient on *MF * POST* toward zero.

Control Variables

Theory is insufficiently rich to provide much guidance concerning the control variables in a system of equations where the dependent variables are investor composition and the mix of

¹⁶ Brown et al. (2007) report that insiders were particularly influential among dividend initiators in 2003. They present evidence consistent with dividends crowding out repurchases in firms with large insider holdings. However, they find no such substitution or insider influence among companies that were paying dividends before JGTRRA, which is the group of firms that is the focus of our analysis.

dividends and repurchases. To our knowledge, no study to date models the non-tax variables that should vary with the dependent variables in this study. Thus, we control for a host of factors that have been found to be associated with the investor mix and the distribution mix.

For the three regression equations testing for clientele effects, we rely on Bushee (2001), who shows that the level of institutional ownership is associated with firm value, and Del Guercio (1996), who documents that institutional holders tend to hold investments that are more prudent. Hence, we include a number of control variables in the investor holdings regressions to capture firm value and the relative quality of the investment. Specifically, we include *SALESGR* as a proxy for firm growth, defined as the average sales growth over the three previous years. We include two proxies for firm risk: beta (*BETA*), which is included to control for systematic risk, and the standard deviation of the prior year's daily market model residuals (*IRISK*) to control for idiosyncratic risk. Market-adjusted returns over the prior year (*MKTADJRET*) is intended to control for firm performance, which has been found to be positively associated with institutional holdings. The S&P common stock rating (*S&PRATING*) and the number of years that the firm is covered by CRSP (*AGE*) are included to capture the relative quality of the underlying investment. The log of the prior year's average monthly volume divided by shares outstanding (*LIQUIDITY*) is included as a control for liquidity because institutional holders prefer more liquid securities. We also include whether a firm is listed on the S&P 500 (*SP500*) as a control because many index funds are required to hold these firms. Finally, we include R&D intensity (*R&DINT*), measured as research and development expenses divided by sales because Hessel and Norman (1992) report that some institutions are fixated on the R&D activity of the firm.

Concerning the fourth regression, where the dependent variable is *DIV%*, we include various measures that are known to affect either dividends or repurchases, although sensitivity tests show that results are largely robust to the set of control variables. First, we include lagged retained earnings scaled by lagged total assets (*RE*). A firm must have earnings and profits (as defined in the tax law) for its distributions to be taxed as dividends. Unfortunately, earnings and profits are unobservable, found only in confidential corporate tax returns. Thus, we use retained earnings as a proxy for earnings and profits. If firms with low or no retained earnings have fewer distributions that qualify as dividends, then *DIV%* should increase in *RE*. Consistent with this expectation and liquidity constraints, DeAngelo et al. (2006) report that firms with low or no retained earnings pay fewer dividends. Next, we include earnings in the model. Jagannathan et al. (2000) and Guay and Harford (2000) report that dividends are paid from permanent earnings, whereas repurchases are paid from transitory earnings. Dittmar and Dittmar (2004) contend that both are paid from permanent earnings, but agree that repurchases come from transitory earnings. Thus, we dichotomize earnings into a permanent component (*PERM*) and a transitory component (*TRANS*). We measure *PERM* with operating income and *TRANS* as the difference between net income and operating income, scaling both components by lagged assets.¹⁷ Based on the conflicts in prior research, we make no prediction about the sign of *PERM*. However, we expect *DIV%* to decrease in *TRANS*. We also include a measure of the firm's payout capacity, free cash flow scaled by lagged assets (*FCF*). Dividends serve as a possible mechanism to reduce agency problems in firms with free cash flow (Jensen and Meckling 1976). Thus, we anticipate that *DIV%* is increasing in *FCF*. We include the lagged ratio of dividends to the market value of equity (*DYIELD*) to control for a firm's capacity to increase dividends. The higher the existing dividend yield, the less the firm will be able to increase its dividend in response to the passage of the JGTRRA. We expect *DIV%* to be increasing in *DYIELD*. We add lagged long-term debt, scaled by lagged assets (*LEVERAGE*), to

¹⁷ Differences between permanent and transitory earnings include special items, other income, and discontinued operations.

control for cross-firm variation in capital structure. Finally, we include the natural logarithm of total assets (*SIZE*) to control for any size effects. We have no expectations about the sign of the *LEVERAGE* and *SIZE* coefficients.

IV. RESULTS

Sample Selection

We begin our tests for investor and manager responses following JGTRRA by drawing an initial sample from the 14,122 firms in the Compustat database between the second quarter of 2001 and the quarter preceding the one that includes May 2003. We then exclude (1) firms whose shares were not common or publicly traded, (2) firms that changed their fiscal year-end during our sample period, (3) financial institutions and insurance companies because regulatory constraints may inhibit management from altering the firm's payout policy, (4) firms with missing Compustat information, and (5) firms not in existence at any time from July 1, 2001, to June 30, 2005.

From the remaining 1,923 firms, we draw two balanced panels. Each firm must have one observation for the eight quarters preceding JGTRRA and one observation for the eight quarters following JGTRRA. Because *DIV%* is undefined if there is neither a dividend nor a repurchase, each firm must have at least one dividend or one repurchase, both before and after JGTRRA. The first sample (Dividend Payers Sample) is the 421 firms in the study that paid dividends sometime during the eight quarters immediately preceding JGTRRA and had either a dividend or repurchase (or both) during the eight quarters immediately following JGTRRA. The second sample (Dividend Payers and Repurchasers) is the 294 firms that both paid dividends and repurchased shares sometime during the eight quarters immediately preceding JGTRRA and had either a dividend or a repurchase (or both) during the eight quarters immediately following JGTRRA.

Our definition of *DIV%* forces us to exclude firms that distribute no profits to shareholders. We further limit our tests to firms that paid dividends at least once during the eight quarters preceding JGTRRA. The reasons for these limitations are twofold. First, firms that were paying dividends before passage paid 97 percent of the dividends issued in the four quarters following enactment. However, much of the JGTRRA research (e.g., Chetty and Saez 2005; Brown et al. 2007) has focused on the relatively narrow impact of JGTRRA on dividend initiation in 2003 alone.¹⁸ Thus, we focus on dividend-paying firms as an under-studied, but more economically significant, portion of the economy. Second, dividend initiation and, to a lesser extent, a firm's first share repurchase convey more and different information to the markets than simply altering the amount of an ongoing stream of dividends or repurchases. Thus, firms that have a history of shareholder distributions can likely modify the dividend-repurchase mix at a lower cost than firms that have never paid dividends or repurchased shares. In fact, it is possible that some managers that had never paid dividends before JGTRRA considered initiating payouts in response to the changed tax incentives associated with JGTRRA, but decided that the costs of initiation exceeded the benefits of attracting individual investors who were now seeking more dividend income. By limiting our analysis to dividend-paying firms, we ensure that the potential costs of dividend initiation do not affect our estimates of managerial responsiveness to the changed individual tax incentives under JGTRRA. This both increases the power of our tests and removes an additional factor for which we would need to control if we included non-dividend-paying firms. One downside to limiting the sample to firms that were already paying dividends is that, if these firms were at their dividend capacity when JGTRRA was enacted, then they may have been unable to increase

¹⁸ This is not to imply that initiations arising from the JGTRRA were unimportant. Initiators began issuing dividends that they might not have paid and likely will continue to issue dividends in the future because dividend payments are sticky and the market takes a dim view of dividend cuts.

their dividend payouts, even if they had wished to respond to the changed tax incentives for individual investors.

Another distinguishing factor about our analysis of JGTRRA is that we examine investor and firm responses through 2005. Desai and Dharmapala (2011) study portfolio rebalancing in 2003 only. Brown et al. (2007) and Aboody and Kasznik (2008) study changes in the dividend-repurchase mix for only two quarters after passage of the legislation. As Shevlin (2008) notes, examining such a short period requires quick response by firms and raises concerns about whether these studies miss important responses in 2004 and 2005. It seems unlikely that investors completely unwound their positions and firms had fully adjusted their notoriously sticky dividends and repurchase policy in response to the tax rate reductions during the seven months following enactment. By investigating a longer window, we permit a more deliberate response and can calibrate how long it took for the largest dividend tax rate reduction in history to fully permeate the economy.

Although we exclude non-dividend-paying firms from our primary tests, in the process of selecting the sample firms, we detect some initial evidence consistent with firms shifting from repurchases to dividends. Among the 1,923 firms from which we draw our samples, we find that 145 companies initiated dividends after JGTRRA, while only 30 firms omitted dividends (a net increase of 115 dividend issuers). Meanwhile, 222 firms began repurchasing after enactment while 370 companies stopped repurchasing (a net reduction of 148 repurchasers). We also find that 26 firms both initiated dividends and ceased repurchasing after passage of the JGTRRA, while only three firms omitted dividends and began repurchasing. Furthermore, among 702 firms that repurchased both before and after JGTRRA, 90 initiated dividends while only 10 firms omitted dividends, a net increase of 80 dividend issuers. On the other hand, among the 408 firms that paid dividends both before and after JGTRRA, 46 began repurchasing after passage, but 58 stopped buying back shares, a net decrease of 12 repurchasers. All of these comparisons are consistent with firms shifting from repurchases to dividends following enactment of the JGTRRA.

Descriptive Statistics

Table 1 provides means and medians for the regression variables, both before and after enactment of JGTRRA, for both samples and reports whether the pre- and post-enactment means and medians are significantly different.¹⁹ The samples provide some evidence that dividends increased following JGTRRA (median total dividends and mean dividends per assets increased in both samples). However, there is less evidence that *DIV%* (i.e., dividends as a percentage of total shareholder distributions) increased after passage (only the difference in medians for the Dividend Payers and Repurchasers sample is significant). Recall, however, that we do not hypothesize about the overall impact of JGTRRA on dividends, repurchases, or *DIV%*. Rather, we predict a more positive association between *DIV%* and individual stock ownership following passage of JGTRRA.²⁰

Both samples show decreases in non-executive individual holdings (*NONEXEC*) and in-

¹⁹ Means are tested using a t-test of the means. The p-values for the medians are the larger p-values using the Wilcoxon and Kruskal-Willis tests.

²⁰ Special dividends are included in *DIV%*. Although the frequency of special dividends increased after the enactment of the JGTRRA, the relative magnitude of special dividends continued to be small as compared to regular dividends and repurchases. For our sample of 421 dividend payers and 294 dividend payers and repurchasers, special dividends constituted only 1 percent (0.4 percent) of aggregate payout made after (before) enactment of the JGTRRA.

TABLE 1
Descriptive Statistics
Means and Medians for Regression Variables

	Dividends 421 Firms (at least one dividend before enactment of the JGTRRA)				Dividend Payers and Repurchasers Sample 294 Firms (at least one dividend and one repurchase before enactment of the JGTRRA)			
	Means		Medians		Means		Medians	
	Before	After	Before	After	Before	After	Before	After
Total Dividends	152	200	17	27*	193	252	19	30*
Total Repurchases	140	251	2	3	200	327	13	9*
Dividends per Assets	0.03	0.04*	0.02	0.03	0.03	0.04*	0.02	0.03
Repurchases per Assets	0.03	0.04	0.00	0.00	0.05	0.05	0.01	0.01**
DIV%	0.78	0.78	0.89	0.89	0.68	0.72	0.75	0.81**
INSIDER	0.10	0.09	0.04	0.04	0.11	0.09	0.05	0.04
NONEXEC	0.35	0.30**	0.32	0.25**	0.35	0.30**	0.30	0.25**
MF	0.27	0.29*	0.26	0.29**	0.26	0.29*	0.26	0.29*
RE	0.32	0.32	0.32	0.32	0.34	0.33	0.34	0.36
PERM	0.16	0.18*	0.13	0.16*	0.18	0.20	0.16	0.18
TRANS	-0.07	-0.07	-0.06	-0.06	-0.08	-0.08	-0.07	-0.07
FCF	0.06	0.02**	0.08	0.03**	0.07	0.03**	0.08	0.04**
DYIELD	0.53	0.44	0.40	0.36	0.55	0.43	0.38	0.37
LEVERAGE	0.22	0.20*	0.22	0.18*	0.21	0.18	0.20	0.17
SIZE	6.85	7.03	6.74	6.93	6.94	7.12	6.74	6.94
S&P RATING	12	13**	16	16**	11	13**	16	16*
AGE	25	25	22	22	25	25	22	22
SP500	0.23	0.24	0	0	0.25	0.26	0	0
LIQUIDITY	-0.50	-0.21**	-0.36	-0.14**	-0.48	-0.23**	-0.30	-0.14**
BETA	0.63	0.92**	0.60	0.92**	0.65	0.91**	0.61	0.90**
IRISK	2.46	1.83**	2.29	1.67**	2.42	2.26**	1.81	1.63**
MKTADJRET	0.12	0.05**	0.12	0.03**	0.11	0.05**	0.11	0.03**

(continued on next page)

TABLE 1 (continued)

	Dividends 421 Firms (at least one dividend before enactment of the JGTRRA)				Dividend Payers and Repurchasers Sample 294 Firms (at least one dividend and one repurchase before enactment of the JGTRRA)			
	Means		Medians		Means		Medians	
	Before	After	Before	After	Before	After	Before	After
SALESGR	0.12	0.26**	0.06	0.19**	0.11	0.23**	0.06	0.18**
R&DINT	0.01	0.01	0	0	0.01	0.01	0	0

* ** Indicates significance at the 5 percent and 1 percent levels, respectively, using a two-tailed test.

This table reports descriptive statistics for the 421 (294) firms that paid dividends (both dividends and repurchases) in the period surrounding the JGTRRA. Total Dividends is dividends as reported by Compustat (data16 $\times$ data61). If the firm uses the treasury stock method, then Total Repurchases is the change in treasury stock (data98). If there is a net decrease in treasury stock, then Total Repurchases is set to 0 (i.e., if there are net stock issuances). If the firm does not use the treasury stock method, then Total Repurchases is the repurchase amount from the statement of cash flows (data93) less decreases in preferred stock (change in data55 and data71). Dividends per Assets is Total Dividends scaled by lagged assets (data44). Repurchases per Assets is Total Repurchases scaled by lagged assets (data44).

Variable Definitions:

DIV% = Total Dividends over the sum of Total Dividends and Total Repurchases;

INSIDER = percentage of shares held by insiders as measured by holdings of directors and officers as reported in Thomson Financial's Insider Filing Data;

NONEXEC = 1 less the percentage of shares that are held by institutional investors as reported in 13-F filings and collected by Thomson Financial's Institutional Holdings database less *INSIDER* less non-officer/director beneficial owners as reported in Thomson Financial's Insider Filing data;

MF = percentage of the firm owned by mutual funds as reported in 13-F filings and collected by Thomson Financial's Institutional Holdings database;

RE = lagged retained earnings scaled by lagged assets (data58/data44);

PERM = operating income (data21 - data5 - data22) scaled by lagged assets (data44);

TRANS = net income scaled by lagged assets (data69/data44) less *PERM*;

FCF = income before extraordinary items (data8) plus interest expense (data22) less the change in the applicable balance sheet accounts: (assets (data44), liabilities (data54), debt (data45, data51)) scaled by lagged assets (data44);

DYIELD = lagged ratio of dividends per share to price, expressed in percentages (data16/data14);

LEVERAGE = lagged long-term debt (data51 + data45) scaled by lagged assets (data44);

SIZE = natural log of assets (data44);

S&P RATING = S&P Common Stock Ranking (SPCSR 7 = A + etc.) + 100;

AGE = number of years the firm is reported on CRSP + 100 as of 2003;

SP500 = 1 if the firm is in the S&P500 Index, 0 otherwise;

LIQUIDITY = log of average monthly volume over shares outstanding for the prior year;

(continued on next page)

TABLE 1 (continued)

$BETA$ = market beta estimated over the prior 12 months;
 $IRISK$ = unsystematic risk, which is estimated as the standard deviation of daily market model residuals over the prior year multiplied by 100;
 $MKTADJRET$ = market adjusted returns over the prior year expressed as a percentage;
 $SALESGR$ = average sales growth over the prior two years ($data2/data4$); and
 $R\&DINT$ = R&D intensity estimated as R&D expense over lagged assets ($data4/data44$); missing R&D is set to 0.



creases in mutual funds holdings (*MF*).²¹ Among the control variables, free cash flow (*FCF*), volume (*LIQUIDITY*), idiosyncratic risk (*IRISK*), and returns (*MKTADJRET*) fell after enactment for both samples. Stock ratings (*S&PRATING*), beta (*BETA*), and sales growth (*SALESGR*) increased after passage for both samples.

Preliminary Regression Results

To provide some perspective for the simultaneous equation results that we report shortly, Tables 2 and 3 show separate OLS regression results for each of the four equations. Table 2 shows selected coefficient estimates from the first three equations, where investor ownership percentages are the dependent variables and *DIV% * POST* is the variable of interest. Table 3 presents results

TABLE 2
OLS
Changes in Ownership by Insiders, Non-Executives, and Mutual Funds around JGTRRA Enactment

	Equation (1): <i>INSIDER</i> (Insider % Ownership)		Equation (2): <i>NONEXEC</i> (Non-Executive % Ownership)		Equation (3): <i>MF</i> (Mutual Fund % Ownership)		
	Pred.	Dividend Payers and Repurchasers	Dividend Payers	Dividend Payers and Repurchasers	Dividend Payers	Dividend Payers and Repurchasers	
Intercept		-0.027	-0.040	0.273**	0.300**	0.414**	0.422**
<i>POST</i>		-0.013	-0.006	0.049	0.022	-0.020	-0.020
<i>DIV%</i>		-0.006	0.021	-0.008	-0.039	0.022	0.003
<i>DIV% * POST</i> (+)		0.048	0.038	-0.004	0.003	-0.026	-0.011
<i>S&PRATING</i>		-0.029	-0.018	-0.144	-0.020	0.002*	0.184*
<i>AGE</i>		-0.084**	-0.063	0.064	0.072	-0.037	-0.076*
<i>SP500</i>		-0.001	-0.011	0.030	0.023	-0.093**	-0.082**
<i>LIQUIDITY</i>		-0.045**	-0.044**	-0.107**	-0.108**	0.094**	0.089**
<i>BETA</i>		-0.011	-0.003	-0.082**	-0.068**	0.030**	0.028**
<i>IRISK</i>		2.211**	2.543**	5.831**	2.968**	-3.785**	-2.783**
<i>MKTADJRET</i>		11.088*	15.281**	14.383	12.398	-7.804	-9.263
<i>SALESGR</i>		0.015	0.019	0.004	-0.019	-0.016	-0.004
<i>R&DINT</i>		0.133	0.191	-0.173	-0.423*	-0.291**	-0.255*
n		842	588	842	588	842	588

*, ** Indicate that the coefficients are significantly different from 0 at the 5 percent and 1 levels, respectively, using a one-tailed test where there are predictions; otherwise, using a two-tailed test.

This table reports estimated ordinary least squares regression statistics for the 421 (294) firms that paid dividends (both dividends and repurchases) in the period surrounding the JGTRRA. The dependent variables are the percentages of ownerships by insiders, non-executives, and mutual funds. Observations aggregate eight quarters before and after enactment of the JGTRRA. *POST* is a categorical variable that equals 1 if the dividend/repurchase was declared in the eight quarters after 5/23/03, 0 otherwise. All other variables are defined in Table 1.

²¹ According to the Investment Company Institute (see <http://www.ICI.org>), from 2001 to 2005 mutual fund ownership of all publicly traded equity securities increased from 21 percent to 25 percent. In addition, over our sample period, individuals increased the proportion of their financial assets held by mutual funds from 40 percent in 2001 to 47 percent in 2005.

TABLE 3
OLS
Changes in Dividends and Repurchases around JGTRRA Enactment

Equation (4): <i>DIV%</i>			
	Pred.	Dividend Payers (n = 842)	Dividend Payers and Repurchasers (n = 588)
Intercept		0.95**	0.94**
<i>POST</i>		-0.03	-0.03
<i>INSIDER</i>		-0.01	-0.02
<i>INSIDER * POST</i>	(+)	0.28*	0.31*
<i>NONEXEC</i>		0.02	-0.04
<i>NONEXEC * POST</i>	(+)	0.02	0.06
<i>MF</i>		0.03	-0.15
<i>MF * POST</i>	(+)	0.04	0.09
<i>RE</i>		0.11**	0.10**
<i>PERM</i>		-0.42**	-0.28**
<i>TRANS</i>		0.14	0.25
<i>FCF</i>		-0.16**	-0.16**
<i>DYIELD</i>		0.11	0.14
<i>LEVERAGE</i>		0.08	0.01
<i>SIZE</i>		-0.02**	-0.02*
Industry Indicators		Yes	Yes

*, ** Indicate that the coefficients are significantly different from 0 at the 5 percent and 1 percent levels, respectively, using a one-tailed test where there are predictions; otherwise, using a two-tailed test.

This table reports estimated ordinary least squares regression statistics for the 421 (294) firms that paid dividends (both dividends and repurchases) in the period surrounding the JGTRRA. The dependent variable is the percentage of shareholder payouts that are dividends. Observations aggregate eight quarters before and after enactment of the JGTRRA. All variables are defined in Tables 1 and 2.

for the fourth equation, where *DIV%* is the dependent variable and the ownership percentages of the three-investor groups are explanatory variables.

Starting with Table 2, we expect a positive coefficient on *DIV% * POST*, which will be interpreted as evidence that, following enactment, individual ownership increased for those firms that distributed larger portions of their profits as dividends. Using OLS, we find a positive coefficient on *DIV% * POST* (α_3) when *INSIDER* is the dependent variable. However, the coefficient is not significantly greater than 0 at the 5 percent level (although using a one-tailed test, the coefficient is significant at the 10 percent level for the Dividend Payers sample). Contrary to expectation, three of the four coefficients on *DIV% * POST* are negative when the dependent variable is *NONEXEC* (δ_3) or *MF* (γ_3), although none are significantly different from 0. In short, the OLS results in Table 2 provide no evidence that investors rebalanced their portfolios following enactment, shifting toward stocks that distributed a larger portion of their payout as dividends.

Among the control variables, two are significant in every regression. *LIQUIDITY* is negative when *INSIDER* or *NONEXEC* is the dependent variable and positive when *MF* is the dependent variable. The signs flip for *IRISK*, as its coefficient is always positive when *INSIDER* or *NON-EXEC* is the dependent variable and always negative when *MF* is the dependent variable. These results are consistent with differences in the non-tax factors that matter to insiders and non-

executive individuals investing on their own account as compared to those that matter to mutual fund investors (see Bushee 2001).

Table 3 presents OLS summary statistics from regressing the ratio of dividends to total payouts (*DIV%*) on the measures of individual ownership and control variables. We predict positive coefficients on *INSIDER * POST* (β_3), *NONEXEC * POST* (β_5), and *MF * POST* (β_7), consistent with firms altering their distribution policy to retain and attract individual investors, given the tax changes in JGTRRA.

We find that *INSIDER * POST* coefficients are positive and significantly greater than 0 at the 5 percent level. We interpret these findings as evidence that firms with large holdings by directors and officers distributed a larger portion of their profits as dividends, after enactment, than they did before enactment. This is consistent with the individuals who set the dividend/repurchase policy (i.e., the directors and officers) modifying the distribution policy after enactment in a manner that is consistent with their own (and other individual) shareholders' interests. The *NONEXEC * POST* and *MF * POST* coefficients also are positive, but not significantly greater than 0 at conventional levels.²² These initial results are consistent with managers responding to the altered tax incentives of directors and officers, but not other individual investors.

Among control variables, the coefficients on *RE* (retained earnings) are always significantly positive. The coefficients on *PERM* (operating income), *FCF* (free cash flow), and *SIZE* (total assets) are always significantly negative.

Primary Regression Results

We now discuss the results from estimating all four regression equations simultaneously. We find that coefficient estimates are larger and more significant using simultaneous estimation procedures, consistent with joint evaluation of investor and firm responses leading to superior estimates than were detected using separate estimations. We infer from the stronger results that investors and managers responded concurrently to the tax changes in JGTRRA.

Table 4 shows the results from estimating all four equations simultaneously using GMM. For brevity, we report only the coefficients for the key variables. Looking first at insider responses, we find that, when *INSIDER* is the dependent variable, the coefficient on *DIV% * POST* (α_3) is significantly greater than 0 for the Dividend Payers sample at the 5 percent level. (It is significant at the 10 percent level using the Dividend Payers and Repurchasers sample.) This finding is consistent with directors and officers boosting their holdings in their own firms, after enactment, if their firms were distributing relatively large portions of their profits as dividends (i.e., it is consistent with a tax clientele effect).²³ The *DIV% * POST* coefficient of 0.11 implies that an increase of one standard deviation in *DIV% * POST*, or 0.28, would boost insider holdings by three percentage points. This is a substantial increase in holdings by directors and officers because they held only 10 percent of the firm, on average, before JGTRRA. In contrast, as in Table 2, we continue to find no evidence that non-executive individuals (δ_3) or mutual fund investors (γ_3) rebalanced their holdings by shifting toward firms distributing a higher portion of their profits as dividends, as compared with repurchases. Thus, we conclude that directors and officers were the only investor group changing their holdings in response to firm distribution policy.

²² The *NONEXEC * POST* coefficient is significant at the 10 percent level using a one-tailed test in the Dividend Payers and Repurchasers sample.

²³ Since the percentage of stock held by insiders was declining during the investigation period (see Table 1), insiders in high-dividend-paying firms may not have been buying far more shares in their own companies. Instead, perhaps these insiders simply sold few shares in their own companies during this period while insiders in lower dividend-paying firms were unloading large holdings.

TABLE 4
GMM Estimation of System of Equations
Changes in Ownership by Insiders, Non-Executives, and Mutual Funds and Changes in
Dividends and Repurchases around JGTRRA Enactment

Equation	Dependent Variable	Explanatory Variable	Pred.	Dividend Payers (n = 842)	Dividend Payers and Repurchasers (n = 588)
(1)	<i>INSIDER</i>	<i>DIV% * POST</i>	(+)	0.11*	0.10
(2)	<i>NONEXEC</i>	<i>DIV% * POST</i>	(+)	0.09	-0.04
(3)	<i>MF</i>	<i>DIV% * POST</i>	(+)	-0.02	0.01
(4)	<i>DIV%</i>	<i>INSIDER * POST</i>	(+)	0.32**	0.42**
(4)	<i>DIV%</i>	<i>NONEXEC * POST</i>	(+)	0.19*	0.18*
(4)	<i>DIV%</i>	<i>MF * POST</i>	(+)	0.28*	0.33*

*, ** Indicate that the coefficients are significantly different from 0 at the 5 percent and 1 percent levels, respectively, using a one-tailed test.

This table reports estimated GMM estimates of the simultaneous equations' regression statistics for the 421 (294) firms that paid dividends (both dividends and repurchases) in the period surrounding the JGTRRA. The dependent variables are the percentages of ownerships by insiders, individuals, and mutual funds and the percentage of shareholder payouts that are dividends. Observations aggregate eight quarters before and after enactment of the JGTRRA. All variables are defined in Tables 1 and 2. Note that all control variables in Tables 2 and 3 are included in the estimation of the model but excluded from the table for brevity.

Looking at manager responses, we find strong evidence that firms held disproportionately by insiders, non-executive individuals, and mutual funds began to distribute a higher percentage of profits through dividends following enactment. When the dependent variable is *DIV%*, the coefficients on *INSIDER * POST* (β_3), *NONEXEC * POST* (β_5) and *MF * POST* (β_7) are all positive and significant at conventional levels for both samples. These results are consistent with managers modifying their payout policy in response to their individual owners' increased preference for dividends compared with repurchases, following passage of JGTRRA. In other words, as dividends became less tax-disadvantaged, firms appear to have provided more dividends as a percentage of their total payout. Brown et al. (2007) report similar findings but only for insiders with large holdings in firms that initiated dividends in 2003.²⁴ Our findings show that firm responses extended to a broader set of firms (non-initiating dividend payers) and individual investors.

Analyzing the economic significance of each coefficient using a one-standard deviation increase, we find that the *INSIDER * POST* coefficient of 0.32 (0.42) with the Dividend Payers (Dividend Payers and Repurchasers) sample implies a four (six) percentage point increase in *DIV%*—i.e., the percentage of payout distributed as dividends. With a pre-JGTRRA mean *DIV%* of 78 percent (68 percent), this suggests that, for every standard deviation increase in insider holdings, the ratio of dividend issuance to total distributions rose 5 percent (9 percent). Likewise, the *NONEXEC * POST* coefficient suggests a four percentage point increase in *DIV%* for both samples. Finally, the *MF * POST* coefficient of 0.28 (0.33) implies a four (five) percentage point increase in *DIV%* for the Dividend Payers (Dividend Payers and Repurchasers) sample. In short, all six coefficients across the two samples imply an economically significant change in distribution policy in response to altered investor preferences following JGTRRA.

²⁴ Brown et al. (2007) treat all 2003 dividend initiations as responses to JGTRRA, even though the legislation was highly controversial and passed by a single vote in the U.S. Senate in May.

We find that the GMM estimates are substantially stronger than those for OLS. This is consistent with our expectation that investor and manager responses are simultaneously determined. Evaluating only investors or only managers potentially explains some of the weak and conflicting results from prior JGTRRA studies (see reviews by Dharmapala [2009] and Shackelford [2009]). If we were to stop with the OLS results, then we would erroneously conclude that no investors rebalanced their portfolios and that the firm responses were limited to those companies with heavy insider ownership. We also would substantially understate the economic significance of legislation as estimated by using the regression coefficients.

Comparisons with Non-Event Periods

This section repeats the analyses detailed above using non-event periods. If the 2003 findings are related to JGTRRA, then we expect the coefficients from the non-event years to be different from the 2003 coefficients. We repeat the tests as if the actual May 2003 rate reductions had occurred in May 1994, and include the eight pseudo “pre-enactment” quarters prior to the quarter that includes May 1994 and the eight pseudo “post-enactment” quarters that follow the quarter that includes May 1994. We then repeat this seven more times, using May of each year as the pseudo event.²⁵

Table 5 reports the GMM summary statistics for the key coefficients for each year for the Dividend Payers sample.²⁶ For example, estimating Equation (1) for 1994, where *INSIDER* is the dependent variable, the *DIV% * POST* coefficient is -0.04 . For Equation (2), when *NONEXEC* is the dependent variable, the *DIV% * POST* coefficient is 0.23 . For Equation (3), when *MF* is the dependent variable, the *DIV% * POST* coefficient is 0.01 . For Equation (4), for the single firm response equation where the dependent variable is *DIV%*, the *INSIDER * POST* coefficient is 0.15 , the *NONEXEC * POST* coefficient is 0.03 , and the *MF * POST* coefficient is 0.02 . Unlike the 2003 results for which four of the six coefficients are significantly positive, none of these six coefficients are significantly different from zero. However, at the same time, none of the 1994 coefficients are significantly different from their 2003 counterparts, suggesting that there may be nothing unique about 2003. Results for the remaining years are tabulated, and summary statistics for the coefficients from the non-event years are presented at the bottom of Table 5.

The findings confirm the inferences drawn from the primary tests in Table 4. The 2003 coefficient on *DIV% * POST*, when *INSIDER* is the dependent variable, exceeds its counterpart in every year except 2001. Using a t-test (sign test), we can reject at the 0.02 (0.07) level the null hypothesis that the 2003 coefficient would be randomly selected from a distribution formed by the eight coefficients from 1994 through 2001. Finding that the 2003 association was different provides further confirmation that insider holdings, following JGTRRA, were increasing in the dividend percentage of total payout. As further confirmation of the Table 4 results, we continue to find no indication of tax clientele movements by other individuals or mutual funds.

Turning to the manager responses to JGTRRA, we find that the 2003 *INSIDER * POST* coefficient always exceeds its counterparts in every year from 1994 to 2001 and significantly so from 1997 to 2000. The 2003 *NONEXEC * POST* coefficient is greater than its counterparts in every year except 1996, and is statistically significantly larger in two of the eight years. The 2003 *MF * POST* coefficient always exceeds its non-event counterparts, though not significantly. However, for all three 2003 coefficients, we can reject at the 1 percent level, using a t-test, that the 2003 coefficient would be randomly drawn from a distribution of the estimated coefficients from

²⁵ We stop with 2001 because any test of 2002 would result in the inclusion of some of the period following enactment of the JGTRRA, invalidating its use as a non-event, comparison period.

²⁶ Inferences are qualitatively unaltered if the Dividend Payers and Repurchasers sample is used.

TABLE 5
GMM Estimation of System of Equations
Changes in Ownership by Insiders, Non-Executives, and Mutual Funds and Changes in
Dividends and Repurchases around Hypothetical, Non-Event Dates

Dependent Variable	Equation (1)		Equation (2)		Equation (3)		Equation (4)		Equation (4)
	INSIDER		NONEXEC		MF		INSIDER * POST		
Explanatory Variable	DIV% * POST		DIV% * POST		DIV% * POST		INSIDER * POST		DIV% * POST
2003 Estimated Regression Coefficients (from Table 4)	0.11*		0.09		-0.02		0.32**		0.19*
Assume Law Had Passed in May of:									
1994	-0.04		0.23		0.01		0.15		0.03
1995	-0.02		0.12		-0.02		0.22*		0.08
1996	0.01		0.21		0.01		0.16		0.20*
1997	-0.02		0.04		0.02		-0.07		-0.02
1998	0.09		-0.14		0.03		-0.18		-0.20*
1999	-0.10		-0.12		0.04		-0.22		-0.15
2000	0.05		0.03		-0.02		-0.17		0.01
2001	0.15*		-0.01		-0.08		0.14		0.01
Descriptive Stats for 1994-2001 Coefficients									
Mean	0.02		0.05		0.00		0.00		-0.01
Median	-0.01		0.04		0.01		0.04		0.01
Standard Deviation	0.08		0.14		0.04		0.18		0.13
Maximum	0.15		0.23		0.04		0.22		0.20
Prob. that 2003 coefficient could be randomly drawn from a distribution of 1994-2001 coefficients using t-test (sign test)	0.02		0.38		0.21		0.01		0.01
	(0.07)		(0.73)		(0.38)		(0.01)		(0.07)
									(0.01)

* ** Indicate that the estimated regression coefficient is significantly different from 0 at the 5 percent and 1 percent level, respectively. Bold coefficients are significantly different from their 2003 counterparts at the 5 percent level.

(continued on next page)

TABLE 5 (continued)

This table reports estimated GMM estimates of the simultaneous equations' regression statistics for the 421 firms that paid dividends in the period surrounding the JGTRRA from Table 4. The dependent variables are the percentages of ownerships by insiders, individuals, and mutual funds and the percentage of shareholder payouts that are dividends. The investigation periods are quarters surrounding various non-event periods. Sample firms issue at least one dividend during the two years preceding the assumed date of legislative passage. *POST* is a categorical variable that equals 1 if the dividend/repurchase was declared in the eight quarters after May 23 of each year, 0 otherwise. All variables are defined in Tables 1 and 2. Note that all control variables in Tables 2 and 3 are included in the estimation of the model but excluded from the table for brevity.

the non-event years. We can similarly reject at the 1 percent (7 percent) level, using the sign test, for the *INSIDER * POST* and *MF * POST* (*NONEXEC * POST*) coefficients. Together, these results support our earlier conclusions that managers of firms with disproportionately large individual investor holdings modified their distribution policy more than did other firms. These sensitivity tests provide comfort about our earlier inferences, providing evidence that the associations in 2003 between distribution policy and individual ownership measures differ from the relations between those variables in the prior decade.

There is a caveat for this robustness check. The capital gains tax rate fell from 28 percent to 20 percent in 1997 without any change in the dividend tax rate. Using the same logic developed in this study, we would have predicted a shift from dividends to repurchases following that rate reduction, albeit of a lesser extent because the rate change was more modest. If that change in the distribution mix did occur, then this sensitivity test using non-event periods would bias in favor of our finding that the 2003 coefficients were more positive than those in previous years. The reason is that the tax incentives facing individual investors after the rate reductions in 1997 would have called for the mix of distributions to shift toward repurchases and away from dividends. In other words, 1997 was not really a non-event year.

Therefore, if firms responded accordingly, then we would expect the coefficients on the individual tax measures for 1997 and perhaps 1998 to be negative. Consistent with firms modifying their payout policy if individuals held disproportionately large interests, the coefficients on *INSIDER * POST*, *NONEXEC * POST*, and *MF * POST* (when *DIV%* is the dependent variable) turn negative for the first time in 1997. Thus, firm responses to changing tax incentives following the 1997 Act may have biased in favor of our rejecting the null hypothesis that the distribution mix did not change after 2003. However, when we exclude the 1997 coefficients in our comparison of the 2003 coefficients to those in the other non-event periods, we still find that the 2003 coefficient estimate remains significantly different from those in the other years. Furthermore, if managers did indeed respond to the 1997 tax cuts by increasing the repurchase portion of their distributions, then these results provide further evidence that managers consider the personal taxes of their shareholders in issuing dividends and repurchasing shares.

How Quickly Did Investors and Managers Respond to JGTRRA?

All of the tests reported to this point aggregate dividends and repurchases over the eight quarters before enactment and the eight quarters after enactment, creating two observations (before and after JGTRRA) for each firm. As discussed above, the reason for aggregation is that, unlike regular, quarterly dividends, repurchases are irregular events, which can lead to highly volatile quarterly measures of *DIV%*. In this section, we relax this restriction and treat each quarter as a different observation, resulting in 16 observations for each firm. The reason we shift to quarterly measures here is to enable us to pinpoint the time when investors and managers responded to JGTRRA.

To get GMM quarterly coefficient estimates for the same key six variables examined throughout this study, we suppress the intercept and include a categorical variable for each quarter from the earliest quarter (eight quarters before enactment—the quarter including May 2001) to the most recent quarter (eight quarters after enactment—the quarter including May 2005). We then interact the categorical variable for each of the 16 quarters with the six variables of interest.

Table 6 reports 16 quarters of coefficient estimates for the six key variables for the Dividend Payers sample. We find little change across the quarters for the three tax clientele regressions. That is not surprising for estimates when *NONEXEC* or *MF* is the dependent variable because we find no evidence anywhere in this study that suggests that non-executive individuals or mutual fund investors rebalanced their portfolios in favor of stocks that pay a larger portion of their payouts through dividends. However, because we find some support for clientele effects among insiders, it

is a bit surprising that we cannot detect any cross-quarter changes in coefficient estimates.

Conversely, for manager responses to JGTRRA, we find a sharp increase in the quarterly coefficient estimates during the second quarter following passage, which would be the quarter ending December 31, 2003, for most companies (i.e., those with March, June, September, and December year-ends). In eight of the nine preceding quarters, including the quarter immediately following the quarter of enactment (quarter +1), the *INSIDER* coefficient is negative. Beginning with quarter +2, the *INSIDER* coefficient is positive in all but one quarter, and is significantly positive in four of the seven quarters. Likewise, the quarterly pattern for the *NONEXEC* coefficients shows a similar break in quarter +2. The *NONEXEC* coefficient is negative for all nine quarters before quarter +2. The coefficient turns positive in that quarter and remains positive in four of the remaining six quarters. Similarly, the *MF* coefficient is negative in six of the eight pre-enactment quarters and negative only once thereafter. The *MF* coefficient is actually positive in quarter +1, but its coefficient triples from quarter +1 to quarter +2.

We conclude from the quarterly findings that companies with disproportionately large individual ownership began to substantially adjust their payout policy in response to JGTRRA during the last quarter of 2003 and the changes were sustained for the remainder of the investigation period. A delay of one quarter before firms responded to the tax incentives of individual investors is consistent with firms being unable or unwilling to respond to the new tax rates in the quarter immediately following passage. Perhaps they had already made their dividend and repurchase decisions for the third quarter of 2003 by May 2003. However, by the last quarter of 2003, it appears that firms held disproportionately by individuals were beginning to shift at the margin from repurchases to dividends.

This delay of one quarter may partially explain Brown et al.'s (2007) inability to find non-dividend initiators substituting dividends for share repurchases in 2003 and Aboody and Kasznik's (2008) failure to link repurchases, stock options, and individual ownership. Both studies treat all dividends paid and shares repurchased in 2003 as post-enactment payouts.²⁷ If few firms were adjusting their distribution policy before the last quarter of 2003 and, in fact, firms were making distribution choices before then, as though prior tax law applied (as implied by the negative coefficients), then it is understandable that their tests would have struggled to detect any movement from repurchases to dividends.

V. CLOSING REMARKS

This study extends our understanding of the effects of shareholder taxes on firm payout policy by estimating a system of equations to quantify the investor and managerial responses to the unprecedented 2003 cuts in dividend and capital gains tax rates. We hypothesize that, in response to the legislation, individual investors (the only ones affected by the rate reductions) rebalanced their portfolios to increase their dividend income, while firms boosted the portion of their profits that they returned as dividends. Comparing firm-level individual ownership and dividend-repurchase mix, before and after 2003, we find evidence consistent with directors and officers

²⁷ Our quarterly results, indicating no response to JGTRRA before the second half of 2003 and very little until the last quarter of 2003, raise concerns about the pre/post JGTRRA classifications in Brown et al. (2007), Chetty and Saez (2006, 2005), and Aboody and Kasznik (2008). Those studies treat all dividends paid and shares repurchased in 2003 as responses to JGTRRA, even those declared months before the May 23rd passage of the legislation, (e.g., Microsoft's initial dividend announcement on January 7, 2003). By using the day that the dividends were paid, they even include some dividends declared in 2002, well before President Bush ever mentioned possible dividend tax relief in January 2003. Brav et al. (2008) state that it is implausible that firms were so clairvoyant that they declared tax-motivated dividends months before passage. Among other factors, President Bush's initial comments were vague and preliminary, and weeks passed before details of his proposal emerged. Furthermore, passage of the highly controversial legislation was uncertain until Vice-President Cheney cast a tie-breaking vote in the U.S. Senate to gain passage of the legislation.

rebalancing their portfolios, but not other individual investors. We also find evidence consistent with those firms with large individual ownership boosting the dividend portion of their total payouts. The payout adjustments began a few months following enactment of the tax cuts. The regression coefficient estimates imply that both the portfolio rebalancing and distribution policy changes following JGTRRA are economically significant.

To our knowledge, this is the first study to evaluate investor and managerial responses to JGTRRA using simultaneous equations. We find that the results are much stronger than they would have been had we used standard OLS estimates of separate regressions, consistent with investors and managers acting concurrently. Both statistical and economic significance increased substantially using GMM. Furthermore, this is the most comprehensive analysis of the JGTRRA to date. Other studies focus on either investor responses or managerial responses; we study both. We analyze the set of firms that pay almost all dividends (i.e., non-dividend initiators), three different groups of individual investors (insiders, non-executive individuals, and mutual funds), and the 16 quarters centered on passage of the legislation. The methodological and sample enhancements enable us to construct a more powerful test of the association between shareholder taxes and corporate distribution policy.

Finally, our results suggest that additional research is needed to understand the role of insiders in the interaction of shareholder taxes and distribution policies. We find that insiders are the only individuals who rebalanced their portfolio in response to the rate changes. We also find that firms were particularly responsive to the changed tax incentives if directors and officers held large positions. These results complement Brown et al.'s (2007) finding that insiders played key roles in initiating dividends at the expense of share repurchases in 2003, which suggests that corporate governance needs to be introduced into analyses of the shareholder tax-payout choice. Questions that future research could explore include the following: To what extent are insiders motivated by their own personal tax considerations when setting distribution policy? How do the tax incentives of insiders affect the returns to other investors? From a distribution policy perspective, did other individual investors benefit from holding stocks in companies with high insider interests while non-individual shareholdings suffered? To what extent does compensation of executives with stock affect a firm's distribution policy? To what extent do changes in shareholder taxes affect a firm's value differently, depending on the extent of its insider holdings? We look forward to answers to these and similar questions in future studies.

REFERENCES

- Aboody, D., and R. Kasznik. 2008. Executive stock-based compensation and firms' cash payout: The role of shareholders' tax-related payout preferences. *Review of Accounting Studies* 13 (2–3): 216–251.
- Auerbach, A. J., and K. A. Hassett. 2007. The 2003 dividend tax cuts and the value of the firm: An event study. In *Taxing Corporate Income in the 21st Century*, edited by A. J. Auerbach, J. R. Hines, Jr., and J. Slemrod, 93–126. London, U.K.: Cambridge University Press.
- Ayers, B. C., B. Cloyd, and J. Robinson. 2002. The effect of shareholder-level dividend taxes on stock prices: Evidence from the Revenue Reconciliation Act of 1993. *The Accounting Review* 77 (4): 933–947.
- Ayers, B., C. Lefanowicz, and J. Robinson. 2004. The effect of shareholder-level capital gains taxes on the structure of corporate acquisitions. *The Accounting Review* 79 (4): 859–887.
- Bagwell, L. S., and J. B. Shoven. 1989. Cash distributions to shareholders. *The Journal of Economic Perspectives* 3 (Summer): 129–149.
- Berlheim, B. D. 1991. Tax policy and the dividend puzzle. *The RAND Journal of Economics* 22 (4): 455–476.

- Blouin, J. 2009. Discussion of: Dividend tax clienteles: Evidence from tax law changes. *The Journal of the American Taxation Association* 31 (1): 23–28.
- , and L. Krull. 2009. Bringing it home: A study of the incentives surrounding the repatriation of foreign earnings under the American Jobs Creation Act of 2004. *Journal of Accounting Research* 47 (4): 1027–1059.
- , J. Raedy, and D. Shackelford. 2003. Capital gains taxes and equity trading: Empirical evidence. *Journal of Accounting Research* 41 (4): 611–651.
- Boudoukh, J., R. Michaely, M. Richardson, and M. Roberts. 2007. On the importance of measuring payout yield: Implications for empirical asset pricing. *The Journal of Finance* 62 (2): 877–915.
- Brav, A., J. Graham, C. Harvey, and R. Michaely. 2008. Managerial response to the May 2003 dividend tax cut. *Financial Management* 37 (4): 611–624.
- Brown, J., N. Liang, and S. Weisbenner. 2007. Executive financial incentives and payout policy: Firm responses to the 2003 dividend tax cut. *The Journal of Finance* 62 (4): 1935–1965.
- Bushee, B. J. 2001. Do institutional investors prefer near-term earnings over long-run value? *Contemporary Accounting Research* 18 (2): 207–246.
- Chetty, R., and E. Saez. 2010. Dividend and corporate taxation in an agency model of the firm. *American Economic Journal: Economic Policy* 2 (3): 1–31.
- , and ———. 2005. Dividend taxes and corporate behavior: Evidence from the 2003 dividend tax cut. *The Quarterly Journal of Economics* 120 (3): 791–833.
- , and ———. 2006. The effects of the 2003 dividend tax cut on corporate behavior: Interpreting the evidence. *American Economic Review Papers and Proceedings* 96 (2): 124–129.
- DeAngelo, H., L. DeAngelo, and R. Stulz. 2006. Dividend policy and the earned/contributed capital mix: A test of the lifecycle theory. *Journal of Financial Economics* 81 (2): 227–254.
- Del Guercio, D. 1996. The distorting effect of the prudent-man laws on institutional equity investments. *Journal of Financial Economics* 40 (1): 31–62.
- Desai, M., and D. Dharmapala. 2011. Dividend taxes and international portfolio choice. *The Review of Economics and Statistics* 93 (1): 266–284.
- Dhaliwal, D., M. Erickson, and R. Trezevant. 1999. The effect of dividend initiations on institutional holdings: A test of the tax-clientele hypothesis. *National Tax Journal* 52 (2): 179–194.
- , and O. H. Li. 2006. Investor tax heterogeneity and ex-dividend day trading volume. *The Journal of Finance* 61 (1): 463–490.
- , ———, and R. Trezevant. 2003. Is a dividend tax penalty incorporated into the return on a firm's common stock? *Journal of Accounting and Economics* 35 (2): 155–178.
- Dharmapala, D. 2009. The impact of taxes on dividends and corporate financial policy: Lessons from the 2000s. In *Tax Policy: Lessons from the 2000s*, edited by A. D. Viard, 199–233. Washington, D.C.: The AEI Press.
- Dittmar, A. K., and R. F. Dittmar. 2004. Stock repurchase waves: An explanation of the trends in aggregate corporate payout policy. Working paper, University of Michigan.
- Fama, E., and K. French. 2001. Disappearing dividends: Changing firm characteristics or lower propensity to pay. *Journal of Financial Economics* 60 (1): 3–43.
- Feldstein, M., and J. Green. 1983. Why do companies pay dividends? *The American Economic Review* 73 (1): 17–30.
- Gordon, R., and M. Dietz. 2008. Dividends and taxes. In *Institutional Foundations of Public Finance: Economic and Legal Perspectives*, edited by A. Auerbach, and D. Shaviro, 204–224. Boston, MA: Harvard University Press.
- Grullon, G., and R. Michaely. 2002. Dividends, share repurchases and the substitution hypothesis. *The Journal of Finance* 57 (4): 1649–1684.
- Guay, W., and J. Harford. 2000. The cash-flow permanence and information content of dividend increases versus repurchases. *Journal of Financial Economics* 57 (3): 385–415.
- Guenther, D., and R. Sansing. 2010. The effect of tax-exempt investors and risk on stock ownership and expected returns. *The Accounting Review* 85 (3): 849–875.
- Hessel, C., and M. Norman. 1992. Financial characteristics of neglected and institutionally held stocks. *Journal of Accounting, Auditing & Finance* 7 (3): 313–330.

- Jagannathan, M., C. P. Stephens, and M. S. Weisbach. 2000. Financial flexibility and the choice between dividends and stock repurchases. *Journal of Financial Economics* 57 (3): 355–384.
- Jensen, M., and W. Meckling. 1976. Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics* 3 (4): 305–360.
- Miller, M., and F. Modigliani. 1961. Dividend policy, growth and the valuation of shares. *The Journal of Business* 34 (4): 411–433.
- Moser, W., and A. Puckett. 2009. Dividend tax clienteles: Evidence from tax law changes. *The Journal of the American Taxation Association* 31 (1): 1–22.
- Shackelford, D. 2009. A response to Dhammika Dharmapala. In *Tax Policy: Lessons from the 2000s*, edited by A. D. Viard, 234–247. Washington, D.C.: The AEI Press.
- Shevlin, T. 2008. Discussion of: Executive stock-based compensation and firm's cash payout: The role of shareholders' tax-related payout preferences. *Review of Accounting Studies* 13 (2–3): 252–265.
- Skinner, D. J. 2008. The evolving relation between earnings, dividends and stock repurchases. *Journal of Financial Economics* 87 (3): 582–609.

How a Systems Perspective Improves Knowledge Acquisition and Performance in Analytical Procedures

Billy E. Brewster

The University of Texas at Arlington

ABSTRACT: Auditors are required to understand dynamic business environments as part of the performance of analytical procedures. Prior evidence, though, suggests that auditors have difficulty understanding such environments. This study reports an experimental investigation of techniques that help auditors to identify incorrect management representations by developing expectations that are both accurate and adaptive to changing business conditions. My predictions suggest that analyzing a dynamic client environment through a systems perspective enhances information-processing ability, which then improves both evidence discrimination and the assimilation of new audit evidence. Results reveal that participants taking a holistic systems-based view of a client environment develop more coherently organized mental models that increase their likelihood of identifying management representations that are inconsistent with industry evidence. Furthermore, these participants more efficiently use their information-processing ability, thereby improving assimilation of newly learned evidence to understand how changing business conditions affect their initial expectations.

Keywords: *analytical procedures; knowledge organization; learning; mental models.*

Data Availability: *Contact the author for data availability.*

I. INTRODUCTION

When performing analytical procedures, auditors should treat any discrepancies between pre-developed expectations and management representations as indicators of heightened misstatement risk (Bell et al. 2005). Auditors develop expectations by understanding how a client's economic environment (entity-level evidence) relates to that client's financial statements (AICPA 2006; Knechel 2007). Because auditors compare their expectations about how financial and nonfinancial measures change over time to management representations of the same,

This paper is based on my dissertation at The University of Illinois at Urbana-Champaign. I thank my committee members, Mark Peecher (Chair), Paul Beck, Bruce Hannon, Joel Pike, and Ira Solomon for their insightful guidance on this project. I also thank Brian Ballou, Dain Donelson, Annie Farrell, Jon Grenier, Pat Hopkins, Steve Kachelmeier, Lisa Koonce, Susan Krusche, Justin Leiby, Lauren Maines, Jeff Miller, Lillian Mills, Drew Reffett, Kenny Reynolds, Sandra Vera-Muñoz, and workshop participants at The University of Illinois at Urbana-Champaign, Indiana University, Miami University, University of Notre Dame, The University of Texas at Arlington, and The University of Texas at Austin.

Editor's note: Accepted by Kathryn Kadous.

*Submitted: October 2008
Accepted: October 2010
Published Online: April 2011*

the complex and evolving nature of today's business environment increases auditors' needs to develop expectations that are not only accurate, but also adaptable to changing business conditions (Bell et al. 1997). The purpose of this study is to investigate techniques for examining entity-level evidence that improve analytical procedures performance in both corroborating management representations and accurately updating expectations when business conditions change.

While current audit approaches and standards require auditors to understand complex business environments, laboratory results suggest that even experienced decision-makers have difficulty learning in, and making optimal decisions about, such settings (Stermann 1989; Diehl and Stermann 1995). Consistently, auditors, particularly inexperienced auditors, have difficulty linking entity-level evidence to management representations (Curtis and Turley 2007), and this inability is often a significant component of audit failure (SEC 2003; Peecher et al. 2007). At least a part of this difficulty is likely due to human cognitive limitations inhibiting the auditor's ability to build accurate and sufficiently detailed mental models that take into account how economic events flow down to affect the client's financial statements (Bell et al. 2002, 2005).¹ Of particular difficulty is understanding the nonlinear movement of entity-level evidence. Since auditors develop expectations by cognitively simulating the time-series behavior of how entity-level evidence relates to the financial statements, these cognitive limitations hinder auditors' ability to develop accurate expectations that can be used as a benchmark for evaluating management representations.

Auditing and management theorists suggest one way to improve performance is to train auditors to adopt a systems-based perspective when analyzing a client's economic environment (Bell et al. 1997; Stermann 2000; Peecher et al. 2007). This systems view involves developing, and organizing, a mental model of the client environment by holistically evaluating causal relationships and how they change over time. Drawing on psychology theory, I predict that organizing a mental model using a systems view will increase information-processing ability, thereby improving the accessibility of auditor's pre-existing knowledge of entity-level evidence. This enhanced processing should lead to more accurate and adaptable nonlinear expectations. Without a systems view, cognitive limitations are expected to force auditors into a reductionist view that considers environment components in isolation and simulates more predictable, and often incorrect, linear expectations.

Prior accounting research provides evidence that using systems-based concepts such as holistic templates (Choy and King 2005) and understanding circular causality (Hecht 2010) can improve decisions. However, these studies do not place the decision-maker in a complex setting that strains cognitive-processing resources and requires using pre-existing knowledge to simulate outcomes for evaluating financial information. Such an examination is needed since practitioners note that understanding and using knowledge from complex business environments is an essential step in solving many ambiguous auditing problems (PwC 2003). Also, no study shows how auditors learn, build, and use systems-consistent mental models of their clients or the cognitive factors and processes that moderate performance improvements.

To test this prediction, I conduct a laboratory experiment in which novices perform an analytical procedures task, requiring them to learn about a dynamically complex client environment. The combination of novices and a complex environment allows me to test the theories underlying this study by placing participants in a setting that would overwhelm their cognitive-processing ability unless they have a well-organized mental model of the client. I prompt participants to analyze the client using either a systems-based perspective, to enable a more coherent organization

¹ Mental models are holistic cognitive representations of a specific phenomenon and are constructed based on an individual's understanding of how pieces of information link and interrelate (Rickheit and Sichelschmidt 1999; Bell et al. 2002). Mental models are comprised of one or many knowledge structures, and the model's level of detail and accuracy is subject to both the individual's understanding of the phenomena and cognitive-processing constraints.

of entity-level evidence, or a reductionist-based perspective, leaving the participants to organize such evidence intuitively. I then manipulate whether management provides an explanation for an account fluctuation that is either relatively more consistent or inconsistent with the entity-level evidence. Finally, to learn how auditors update their mental models for changing conditions, I introduce new entity-level evidence that should change the nature of the auditors' existing expectations.

Consistent with predictions, I find that systems participants provide more coherently organized mental simulations of the environment's causal relations. They perform these improved simulations without increasing their cognitive effort, suggesting that the systems view helps participants overcome the previously mentioned cognitive limitations by enhancing their information-processing ability (Clark et al. 2006). Systems participants are more likely than reductionist participants to incorporate key nonlinear industry attributes into their expectations, and are also more likely to identify management representations that are inconsistent with entity-level evidence. Finally, the systems participants who have coherently organized mental models are better able to more accurately update their expectations when presented with new entity-level evidence.

This study extends the existing literature on improving knowledge acquisition and performance during analytical procedures. Prior studies show that auditors' knowledge acquisition for analytical procedures is improved through feedback or increasing cognitive effort (Bonner and Walker 1994; Earley 2001; Moreno et al. 2007). However, in an audit setting, feedback is not always readily available and increasing cognitive effort is not always effective due to the complexity of some analytical procedures tasks. Consequently, audit firms and educators have explored methods to improve auditor mental models, but these methods expect auditors to make the essential causal linkages and simulations of the client environment that prior laboratory evidence suggests is difficult to do without assistance (Bell et al. 1997; Bell et al. 2002). Indeed, previous studies show that such approaches can actually introduce biases into the judgment process (O'Donnell and Shultz 2005). I demonstrate that taking a systems view improves participants' analytical procedures performance both initially (i.e., when gathering entity-level evidence and analyzing management representations), and when presented with new, incoming evidence.

Furthermore, I provide insight into how auditors build mental models by examining the factors that moderate the relationship between the organization (structure) of auditors' entity-level knowledge and their information-processing ability. This relationship accounts for substantial variation in task performance. While Borthick et al. (2006) find that better-developed knowledge structures improve performance, there are few studies examining how auditors build these structures and use them to form judgments. In addition, theoretical learning frameworks suggest that knowledge structure and processing ability are related (Kyllonen and Woltz 1989; Libby 1995), but to date, there is little empirical evidence of the true nature of this association (Ceci et al. 2003). I find that coherently organized knowledge structures enhance information-processing ability, which then improves the likelihood of accurately updating existing knowledge structures with new entity-level evidence. Understanding such a relationship allows audit firms and educators to tailor instruction and training to exploit such learning improvements.

In Section II, I describe relevant research and develop hypotheses. Section III presents the experimental method used to test the hypotheses, and Section IV discusses the results of the experiment. Section V concludes.

II. THEORY AND HYPOTHESES

Background on Learning and Knowledge Acquisition

Understanding how auditors learn and update expectations using a systems view requires examining the cognitive factors that facilitate learning entity-level evidence. Learning is the acquisition and organization of *knowledge* through *working memory* (Bonner 2007). *Working*

memory, the cognitive system for holding and manipulating information, is the intersection between long-term memory and information received through the senses, and is a significant component of intellectual and problem-solving ability (Baddeley 1998; Conway et al. 2003). *Knowledge* is information that is stored and retrieved from long-term memory (Libby 1995; Bonner 2007), and is comprised of content (specific information pieces) and the organization, or structure, of that content. When forming judgments, individuals retrieve knowledge from long-term memory and manipulate it within working memory.

Humans have difficulty learning and performing complex tasks due to limitations within working memory. Individuals can cognitively process only a few information items at a time, and their ability to hold and manipulate items in working memory rapidly decreases as the items' interactivity increases (Engle and Kane 2004). In other words, information with numerous causal linkages or dependences is difficult to learn because individuals cannot hold both the items and their interactions within working memory. One way to increase working memory processing capacity is by developing schemas, or clusters of information (Clark et al. 2006). Schemas form by connecting pieces of information into a coherent and meaningful structure, sometimes because of considering a single example problem (Ahn et al. 1992). Schemas are held in working memory as a single item. Organizing knowledge in long-term memory into schemas eases the strain on working memory by reducing cognitive load and therefore substantially increases cognitive-processing ability.² Improved schema development leads to better-organized knowledge structures, which are typically associated with improved decision-making performance (Sternberg 1997).

Individuals learn by elaborating, or mentally simulating, how new information relates to existing knowledge in long-term memory (Bonner 2007). Mental simulations are cognitive representations of past or potential future events (Kahneman and Tversky 1982). Individuals simulate both by cognitively manipulating how new or existing variables affect an existing mental model and by examining the manipulation's resulting output (Klein 1998; Sanna 2000). Therefore, facilitating the simulation process through mental model development is an important element of learning. Improving auditors' analytical procedures performance requires enhancing their likelihood of accurately simulating expectations of entity-level evidence based on information in their long-term memory.

Using a Systems Perspective to Improve Auditors' Mental Models

To develop expectations about a client's financial position, auditors should construct a sufficiently detailed and accurate mental model of how a client interacts with its business environment (Bell et al. 1997; Bell et al. 2002, 2005). The types of business environments modeled by auditors are complex systems that include features, such as feedback loops and time delays, that can produce nonlinear movement of key system components (Sternman 2000). For example, the time-series movements of entity-level evidence indicators, such as commodity prices and inventory balances, are the culmination of multiple causal effects having different strengths over time. Making accurate inferences and expectations about such indicators requires that decision-makers have mental models that include a coherent sequence of causal relations allowing the auditor to simulate the indicators' movements.

Decision-makers have difficulty learning in, and modeling, complex business environments with nonlinear behavior even when they are presented with all necessary information (Sternman 1989; Sweeny and Sternman 2000). This difficulty results from working memory limitations that prevent individuals from combining and correctly organizing the necessary causal relations caus-

² Cognitive load is the burden that performing a task places on a decision-maker's cognitive system, and increases by specific aspects of a task, such as complexity (Paas et al. 2003).

ing such behavior (Sterman 1989). Working memory constraints hinder retrieving and using knowledge in long-term memory while performing mental simulations. Consequently, when placed in an overwhelming business setting, decision-makers often resort to a heuristic-driven reductionist perspective that considers the system components in isolation and bases decisions on more predictable linear relationships (Sterman 2000). Reductionist-based thinking is a common cause of judgment error in complex business settings (Reason 1990; Sterman 2000).

Auditors can reduce reliance on reductionist thinking by taking a systems-based perspective of the client environment (Richmond 1994; Peecher et al. 2007). Bell et al. (1997, 14) describes a systems view as one that “emphasizes basic principles of organization—how the parts are inter-related and coordinated into a unified whole.” Such a perspective entails focusing on the causal relationships among system components as well as the nonlinear behavior resulting from these system interactions (Richmond 1994; Sterman 2000).³ Auditors taking a systems perspective should more coherently organize the environment’s causal relations in their mental models, thereby enhancing their ability to develop causal schemas. Therefore, auditors adopting a systems-thinking perspective, and developing schemas of the various environment causal relations, are likely to process more information concurrently in working memory and thereby overcome some of the aforementioned cognitive constraints.⁴ These auditors’ mental model components should show a logical causal sequence that allows them to better understand the dynamic nonlinear features of a client environment. This improved coherent organization allows the auditor to more easily access knowledge from long-term memory and simulate how entity-level evidence affects the client. Because these auditors can better mentally simulate how economic events flow down to affect the client, they will more likely develop accurate expectations of industry dynamics.

H1: Participants using a systems perspective will develop expectations regarding entity-level evidence that are more accurate than will participants using a reductionist perspective.

The following hypothesis examines the cause of these underlying improvements by investigating whether a systems perspective does in fact reduce auditors’ cognitive-processing constraints when performing mental simulations. Prior research shows that prompting auditors to perform simulation-like tasks can improve performance (Earley 2001; Moreno et al. 2007), but this research does not examine how to improve performance when working memory is overwhelmed and increasing cognitive effort is ineffective. As noted earlier, a systems view should help auditors organize information into schemas, thereby expanding working memory capacity. To examine the improvements resulting from expanded working memory, researchers measure the *efficiency* of

³ Systems thinking is a byproduct of system dynamics, which is an interdisciplinary field that uses mathematical computer models to simulate complex quantitative problems that decision-makers cannot perform without such aid (e.g., hurricane forecasting, predicting spread of viruses). System dynamics views the world as a web of interacting complex relationships including feedback loops and time delays. Systems theorists have tried to improve individual decision-making by training people to adopt the same principles that make the system dynamic models successful (i.e., understanding multiple causality, feedback loops, etc.). This training typically involves either building computer simulations or developing skill sets that can recognize dynamic interactions. To date, there is limited laboratory evidence showing either the effectiveness of such techniques or an efficient way to implement such methods.

⁴ A system is a “collection of parts that interact to function as a whole” (Bell et al. 1997, 14). As described in this paper, a systems view understands phenomena in terms of these interacting and interdependent parts. In contrast, a reductionist view understands phenomena as isolated individual components. I consider these views as parts of a continuum with system and reductionist endpoints. As individuals construct mental models to guide their viewpoints, they will fall somewhere on this continuum. Information-processing ability is one factor that moderates placement on this continuum and a lack of such ability pushes an individual more toward a reductionist perspective. In this study, I use the term “systems view” as a viewpoint that is relatively more toward the systems endpoint, and “reductionist view” as one that is relatively more toward the reductionist endpoint.

information processing (Paas et al. 2003). Processing efficiency is a byproduct of working-memory capacity and considers both performance and exerted mental effort concurrently. Information processing is more efficient if increases in task performance are higher than expected considering mental effort level, or if mental effort is lower than expected considering performance.

- H2:** Participants taking a systems perspective will process information in working memory more efficiently when performing mental simulations than will participants taking a reductionist perspective.

Improving Performance: Evaluating Management Representations

During analytical procedures, auditors compare client-provided values, either account balances or other amounts provided in explanations, to quantitative expectations of entity-level evidence (Koonce 1993; Hirst and Koonce 1996). Psychology theory suggests that when a client provides an amount, auditors will simulate a range of potential values to compare against that amount (Mussweiler and Strack 2001; Epley and Gilovich 2006). If the client-provided amount is inconsistent with the auditor's simulated range, then the auditor discounts the representation's credibility. Consequently, accurately simulating this range of potential values increases the likelihood of identifying management representations that are inconsistent with entity-level evidence.

Prior literature examines this expectation violation phenomenon, but these studies either (1) rely on the auditors' prior experience to understand entity-level evidence (Choo and Trotman 1991; Earley 2002), or (2) provide participants with a means to examine *concurrently* entity-level evidence along with client-provided evidence (Earley 2001; Vera-Munoz et al. 2007). To date, no studies separately examine the mental model and expectation development process and how they affect *subsequent* evaluation of client representations. Distinguishing between expectation development and evidence evaluation is important because auditors often rely on their prior knowledge for corroborating client-provided evidence, and extant literature shows that auditors can be influenced if they cannot discredit such evidence (Kennedy 1995; Earley et al. 2008).

The aforementioned theory suggests that an ordinal interaction exists between client analysis perspective and management representation consistency with entity-level evidence. Systems auditors will be better able to mentally simulate a range of potential values for entity evidence and therefore distinguish between management representations that are consistent with such evidence from those that are not. These auditors are more likely to provide lower credibility assessments for inconsistent representations. Furthermore, when management representations are inconsistent with entity evidence, systems auditors should be more likely to recognize the inconsistency and provide lower credibility ratings than reductionist auditors. The following hypothesis sets forth the prediction:

- H3:** Participants will evaluate the credibility of management representations at the lowest level when they use a systems perspective of entity-level evidence and when they receive management representations that are inconsistent with that evidence; participants will evaluate the credibility of management representations at a higher level under all other combinations of analysis perspective and evidence consistency.

Improving Performance: Updating Expectations of Entity-Level Evidence

In addition to improving processing ability, a coherently organized mental model provides a structure for assimilating and retrieving incoming information (Brewer and Nakamura 1984; Sweller et al. 1998). In an audit context, Bonner et al. (1997) find that relevant knowledge

structures in place prior to experiencing events aid in learning from those events. I extend this result by examining one potential component of these learning improvements, namely that organizing information reduces the constraints on working memory capacity and, hence, improves an individual's understanding of how to accurately integrate new information to revise a knowledge structure.⁵ While theoretical frameworks suggest that knowledge structure and working memory interact in the learning process (Kyllonen and Woltz 1989; Bonner 2007), there is little empirical evidence showing how this process occurs. This study tests how organized knowledge improves mental simulation in working memory allowing for improved mental model building.

A better-developed mental model established through a systems-thinking perspective likely improves auditors' learning in a dynamic business setting for two reasons. First, the resulting improved simulation coherence should allow the decision-maker to effectively simulate how the various system components can affect each other in a logical sequence. Second, schemas of the causal interactions will reduce cognitive load and increase the likelihood of understanding how the new information affects the dynamic nonlinearities resulting from component interactions. Understanding the environments' dynamics allows decision-makers to manipulate how the new information affects the essential causal relations within their existing mental models.

Auditors using a better-organized mental model with a systems-based perspective will be more likely to understand how new evidence affects the existing cause-and-effect relationships and resulting nonlinearities within their mental model. This understanding makes these auditors more likely to change their existing inferences regarding such evidence when they receive new, contradictory information (Bell et al. 2005).

- H4:** When participants receive new entity-level evidence, those who have relatively more coherently organized mental models and take a systems perspective will most accurately update their mental models to modify existing expectations regarding entity-level evidence compared to all other participants.

III. METHOD

Participants

The participants are 61 accounting students from various junior- and senior-level accounting courses at a large state university, randomly assigned to experimental conditions. The theories underlying this study apply to both expert and novice auditors, but for present purposes, using novices has distinct advantages. First, using student participants shows that less experienced auditor subjects garner process improvements through systems-thinking, contrary to recent critical claims that only experienced auditors can perform the complex business environment analysis required by professional standards (AICPA 2006; Curtis and Turley 2007).

Second, this subject pool affords a more powerful test of theory than a professional subject pool (Smith 1989; Friedman and Sunder 1994).⁶ Experimental tasks must strain working memory to reveal the differences between conditions (Tindall-Ford et al. 1997). The task I employ is more

⁵ My predictions suggest that improved working memory processing ability is *one component* of mental model updating. I do not preclude that improved knowledge structures would benefit updating even if one held processing ability constant (e.g., an impoverished model might not have developed the links essential to understanding how the new information would affect the existing model). My experimental design addresses this issue by examining whether task-performance improvements are due to knowledge content differences alone and/or include an aspect of processing ability.

⁶ Smith (1989) conjectures that the primary purpose of laboratory experiments is to test theory and researchers should design experimental procedures to maximize this goal. Furthermore, Friedman and Sunder (1994, 21) state that experi-

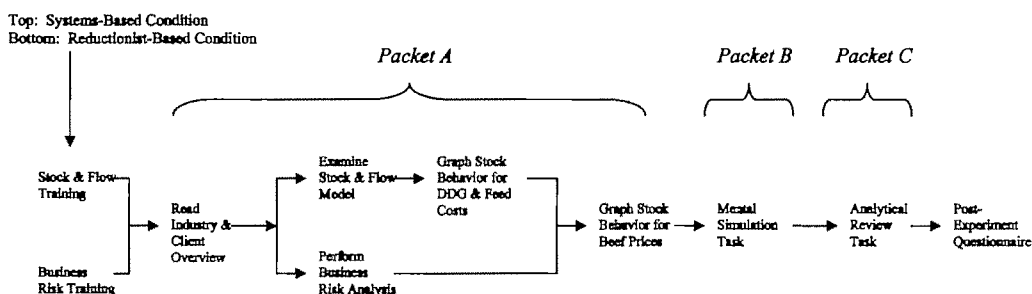
likely to strain working memory, and to do so to a greater degree, for nonprofessional, student subjects than audit professionals because efficient use of working memory resources is a common attribute of expertise (Sternberg 1997).

Also, because my experiment involves training in systems-thinking and the process of learning entity-level evidence, using auditor participants with variable, specialized, pre-developed routines for or disposition toward analyzing such information would not facilitate a clean mental model manipulation. Using student participants with less variable, less specialized, and more malleable knowledge structures, relative to professionals, allows me to effectively control and manipulate mental model development (Libby et al. 2002). Holding constant the participants' pre-existing knowledge structures by using nonprofessional subjects—an approach consistent with previous research examining learning environments (Bonner and Walker 1994; Bonner et al. 1997)—is conducive to identifying how knowledge organization and information-processing ability interact to improve performance.⁷

Task

The instrument contains tasks (see Figure 1) representing analytical procedures performance from initial evidence gathering and mental model building to subsequent management representation evaluation (Koonce 1993; Bell et al. 2005). First, participants receive a tutorial describing how auditors use entity-level information for developing client mental models. Participants then

FIGURE 1
Overview of Experimental Procedures



This figure represents the temporal sequence of the experimental materials that the participants received. Each Packet was taken from the participants before they received the additional items (although they were allowed to use the training tutorial while completing Packet A).

mental design should “sharpen the effects of focus variables and minimize blurring due to nuisance variables.” Controlling for pre-existing knowledge by using nonprofessional subjects supports this objective by minimizing any undue effects this knowledge could have on my dependent variables.

⁷ In addition to these theory-testing advantages, this research pool is also sufficient for my study; auditor participants are unnecessary because no current theoretical or empirical evidence suggests that pre-existing knowledge would interact with the manipulated factors in such a way as to change the *nature* of the independent variables' predicted influence on the dependent measures (Peecher and Solomon 2001). While the pre-existing knowledge structures of professional auditors could reduce the magnitude of the predicted effects by decreasing strain on working memory, there is no reason to suspect that these structures could also cause a predicted directional effect to reverse (i.e., a disordinal interaction).

opened Packet A, which contained materials for a fictitious hamburger food chain.⁸ These materials include an industry and company overview characterizing current conditions in the beef industry (see Exhibit 1). The case describes how increased ethanol production causes the price of corn-based feed for cattle to increase, subsequently increasing beef prices. However, another key point in the case is that the use of an ethanol byproduct, called DDG, for cattle feed causes beef producers' feed costs to decline after June of the period under examination because DDG costs only 60 percent of typical corn-based feed. See Figure 2 for a computer simulation of beef price behavior based on the facts in the case. The case states that the company's success ties heavily to the beef industry and instructs participants to develop expectations by inferring the time-series behavior of beef prices for the year under review.

After turning in Packet A, participants applied their existing mental models to perform a mental simulation task in Packet B by writing a paragraph describing how an abrupt change in ethanol prices would affect the company's profits. After turning in Packet B, participants conducted an analytical procedures task in Packet C. This task shows that the company's profitability declined over the past year and provides both a gross profit analysis and an explanation from management targeting rising beef prices as the reason for the decline. The provided explanation is, however, incomplete, as beef prices did not steadily rise during the year, but declined after June because of implementing the DDG byproduct as feed. The participants assessed the credibility of management's explanation of the profit decline and provided their own estimates of year-end beef prices. Finally, participants received an additional paragraph that explained new facts regarding entity-level evidence and provided updated beef-price estimates considering those facts.

Independent Variables

The experiment employs a 2×2 between-subjects design with two manipulated independent variables: (1) *industry analysis perspective* (systems-based or reductionist-based) and, (2) *the degree to which management representations regarding year-end beef prices are consistent with industry evidence* (consistent at \$125, or inconsistent at \$165).

The industry perspective variable provides conditions that aid participants in either developing a systems view or resorting to a reductionist view. The systems-based condition provides training and materials helping participants to develop a mental model with a coherent structure of the causal relations in the case materials. Existing research does not provide a systems-based manipulation that actually improves quantifiable aspects of auditors' mental models. Therefore, my systems manipulation uses stocks and flows, which systems theorists state is the most effective way to develop accurate mental models.⁹ Business settings are comprised of multiple stocks and flows, and these items should help individuals simulate the nonlinear movement of system components. However, even experienced individuals have difficulty understanding these concepts without extensive training (Sterman 2000). Therefore, the systems participants in my study used methods that are difficult to learn without such training. Consequently, because my study utilizes a brief tutorial, I rely on three commonly used techniques by systems trainers to develop mental

⁸ The name of the hamburger food chain used in the experimental materials is suppressed due to the fact, unknown at the time of the experiment, that a real-life hamburger food chain with a similar name exists. That chain operates in a different part of the U.S. than the university where the participants were recruited, so I have no reason to believe that any participants associated the materials with the actual chain.

⁹ Stocks are accumulations within a system, while flows represent the stock increases and decreases over time. To illustrate with a commonly used example, the water in a bathtub would represent a stock, whereas the water coming through the faucet and leaving through the drain are flows (Sterman 2000). A system's dynamic properties (feedback loops, time delays) are a byproduct of the interactions among and behavior of stocks and flows.

EXHIBIT 1**Beef Industry and Company Overview**

- Beef prices have increased in recent years. The price of beef as of January 1, 2006 was \$100 per hundredweight (cwt) of cattle and reached record levels later in the year.
- The recent increase in beef demand is a sharp contrast to prior periods as the demand for beef was on the decline for the previous 20 years. Consumers moved away from beef to alternatives, such as poultry, that are less expensive and considered healthier. For example, in 1975 beef accounted for roughly 49 percent of meat products consumption but by 1997 that share had fallen to 32 percent. This decline is due to the previously mentioned health concerns about beef as well as the increased prices of beef compared to cheaper alternatives such as pork and poultry. In addition, the beef industry has historically lagged behind both the poultry and pork industry in developing pre-packaged, pre-cooked products that allow consumers to reduce cooking time.
- In the past decade, health concerns by customers have made the industry focus on “natural” beef, which is becoming an expectation of a large and growing portion of the population. For example, consumers in some markets are willing to pay as much as 25 percent more for a round roast if it has a label containing the word “natural.”
- Profitability in the beef industry is declining due to rising production costs. Unfortunately, the long life cycle of cattle means that beef producers must make production decisions years in advance in order to respond to market conditions. Cattle have the longest life cycle of any livestock produced, since it takes cattle four years to reach maturity. Furthermore, the litter size for cattle averages only one calf per litter.
- Like all commodity markets, the price of cattle is dependent on supply and demand. In livestock markets, however, production cannot take place without a breeding stock. Producers must choose whether to send their stock to market or to keep animals on the farm to increase the future supply. Producers control the supply of breeding stock, and that stock is increased when producers withhold mature animals from market. Producers will increase their desired breeding stock (to increase future supply) when the expected markup ratio is high and will reduce their breeding stock when the markup is low. In the short term, the longer life cycle of beef reduces the producer’s ability to increase capacity to deal with rising demand. Industry experts believe it takes between three and four years for cattle producers to fully respond to price increases by increasing production capacity.
- Over the previous 100 years, the price of feed was relatively stable and predictable when adjusted for inflation. In the United States, almost 60 percent of corn production is used as animal feed. Recent trends in the corn industry have caused demand and prices for corn to increase. Ethanol production has increased the demand for corn, which in turn has increased feed costs for cattle producers. The price of corn affects beef and dairy production because added costs affect profit margins of beef producers. The ethanol industry’s demand for feed grain will grow exponentially over the next five years increasing from less than 5 percent of supplies a few short years ago to perhaps 75–80 percent in 2009 if the United States produces an 11–12 billion bushel crop. Typically any price changes in feed costs are reflected in beef prices within a week.
- Over the next few years, ethanol prices are expected to continually rise. The price of feed per bushel was \$3.50 on January 1, 2006 and changed throughout the year.
- One consideration of the ethanol process is the production of the Dry Distiller Grains (DDG) byproduct. This byproduct costs approximately 60 percent of what regular corn feed costs and can be fed to cattle. The availability of DDG as a source of feed is dependent on the available supply. Ethanol producers did not start stockpiling DDG until the beginning of January and did not have enough supply to sell until approximately June 2006.
- Another potential effect of high corn prices is that ranchers could bring lighter cattle to market. Livestock weights have been 10 to 20 pounds lower than usual the past year. This will affect the quantity of beef on the market, which will mean less top-quality meat on the market, which will drive up beef costs.

Company Overview for 2006

- XYZ Company is a publicly held, regional fast food restaurant chain specializing in hamburgers. The Company has more than 600 locations and is located in ten different states in the southwestern United States. The corporate headquarters is located in Fort Worth, Texas.

(continued on next page)

EXHIBIT 1 (continued)

- XYZ Company is differentiated from its competitors by serving its hamburgers made-to-order. Each burger is cooked immediately after the customer orders, with a typical cooking time between 5–7 minutes. Management believes this approach ensures a quality product and allows the customers to see the difference in service and quality between XYZ Company and its competitors.
 - The company takes great pride in the low turnover of its employees. The corporate office does an annual employee survey to get feedback from its restaurant associates. Furthermore, past results have shown that those stores with the lowest turnover margin are the ones with the highest sales. Since 2002, the associate turnover rate has dropped by 92 percent.
 - Despite a history of profitability, management is concerned about the rising beef costs. Approximately, 70% of the company's sales are from beef products, which is substantially higher than most of their competitors that introduced chicken products years ago. Management believes these rising costs will reduce profitability, which will hinder their own ability to expand the menu by introducing a line of chicken products.
-

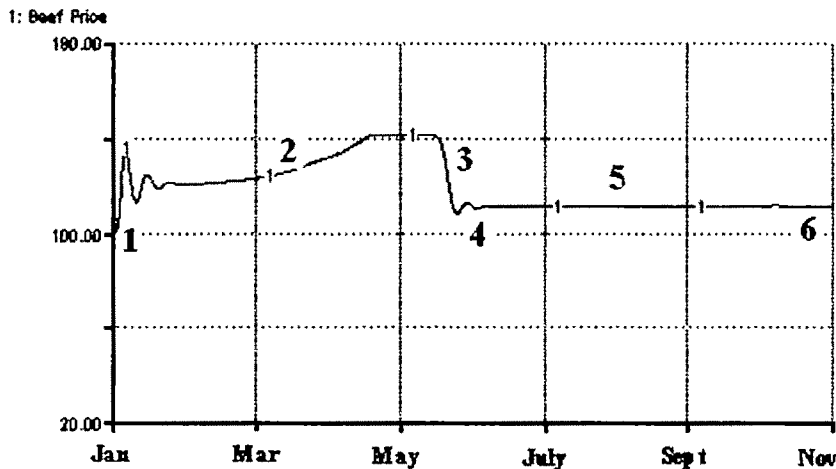
models: (1) diagrams that visually display the causal interactions, (2) active learning techniques to ensure that participants focus on the necessary causal relations, and (3) decomposition of complex causal interactions.

The systems-based participants began the experiment with a tutorial that discusses the nature and interaction of stocks and flows, provides examples, and asks participants to sketch how a stock balance would fluctuate over time. Packet A provides a stock-flow diagram, *based entirely on the industry overview*, showing how the major components of the beef industry interact (see Figure 3). Participants completed an active learning task by writing a sentence describing each interaction within the model. Finally, participants sketched the patterns of behavior for three stocks within the system that have a causal relationship (DDG byproduct, feed prices, and beef prices) using blank graphs similar to that shown in Figure 2.

Participants in the reductionist-based condition used their own intuition to organize a mental model of the case materials. The reductionist participants received the exact same industry and company overview as the systems participants. While they did not have access to the stock/flow tutorial or diagram, I included steps to enhance their understanding of the case materials. First, to ensure that participants understood why constructing a mental model of industry evidence is important and a potentially complex task requiring the assimilation of numerous causal relations, I provided them with a tutorial on business risk auditing, drawing heavily on current auditing standards and methodologies (Bell et al. 1997; Bell et al. 2005; AICPA 2006). Second, I included an active learning task to ensure that the participants studied the industry overview thoroughly enough to construct a mental model of the client environment. Participants identified business risks from the industry overview and described how these risks affect the company. Finally, to reinforce the notion that environment components change over time, participants sketched out the behavior of beef prices over the year.

Both perspective conditions are consistent with current professional standards and business risk auditing methods. The difference in each condition is in the level of aid provided to organize the multiple causal relations in the case. Note that the reductionist participants did not receive the stock-and-flow model, but the information in the industry overview described the facts needed to understand the interactions and nonlinear behavior within the industry *assuming there were no working memory constraints*. The reductionist participants received all the necessary information to understand the nonlinear movement of beef prices. Consequently, I expect them to understand any individual causal relations, but not be able to properly combine them due to working memory

FIGURE 2
Computer Simulation of Beef Prices



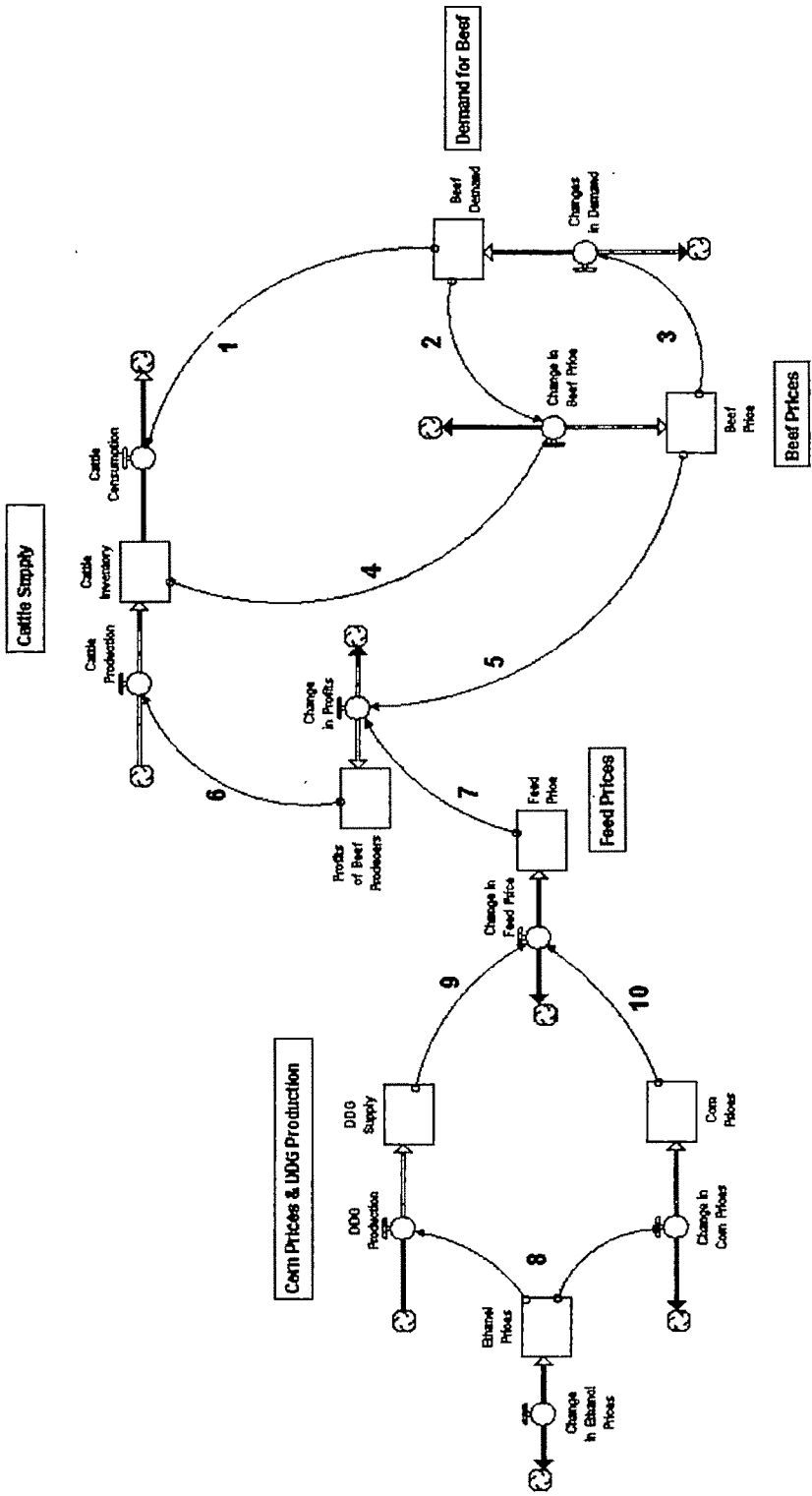
This figure shows a computer simulation of beef prices using the information from the industry overview (Packet A). This simulation model is a supply-and-demand market for commodities based on dynamic models provided in Sterman (2000). The model is coded to move to equilibrium in the absence of external forces as shown by the steady-state periods. The numbers above (1–6) represent features of the simulation used to code the students' graphs of beef price. A point was awarded if the graph exhibited evidence of the above features. The numbers represent the following:

1. The starting point for beef prices is \$100 cwt.
2. Beef prices should rise until approximately June. The rise is due to increasing feed costs as the prices of corn-based feed are rising due to the increased usage of corn for ethanol production.
3. Decline in prices after June as beef producers begin to switch to the ethanol byproduct, DDG, which is 60 percent cheaper than regular corn-based feed.
4. Prices should reach a low point in the periods following June.
5. The market moves back to equilibrium following the decline in feed costs. This can be represented in an increase from the lowest point as long as it is not higher than the pre-DDG period.
6. The final estimate should be lower than the highest point of the graph to reflect the decreased feed costs for beef producers.

limitations. During the post-experiment questionnaire, I verified that participants in both conditions understood the basic causal relations related to beef prices.

My second manipulated independent variable is the consistency level of management representations compared to entity-level evidence. The objective is to provide a representation that is judgmental and requires participants to simulate an outcome from their mental models to evaluate the representation. In the case, management explains that the company's deteriorated profitability stemmed from rising food costs due to increased beef prices. Management states that beef prices had risen to record highs over the past year, and provides a year-end beef price of either \$125 per hundredweight of cattle (cwt) (consistent) or \$165 cwt (inconsistent). Recall, that while participants initially developed expectations, they had a framework (see Figure 2) with a lower boundary

FIGURE 3
Beef Industry Model



This figure represents the model that was provided to students in the Systems-based condition. The participants were instructed to write a sentence describing the effect that each numbered linkage had on the related stocks. The following represents the linkages in the model above. The information below was *not* presented to

the participants, but is included here to help the reader understand the environment:

1. An increase in the demand for beef increases the amount of cattle consumed, and *vice versa*.
2. The demand for beef is one factor that drives beef prices. An increase in demand increases prices, and *vice versa*.
3. Beef prices affect the demand for beef. An increase in prices will decrease demand, and *vice versa*.
4. The supply of cattle directly affects beef prices. An increase in supply will decrease beef prices, and *vice versa*.
5. Beef prices are one component of the profits for beef producers. An increase in price will increase profits, and *vice versa*.
6. The profits of beef producers will affect the amount of cattle produced. As profits rise, there will be more desire to increase production to take advantage of higher profit margins.
7. Feed prices will directly affect the profits of beef producers. An increase in feed prices will reduce profits, and *vice versa*.
8. Ethanol prices affect both the level of DDG byproduct and corn prices. As ethanol prices increase, production will increase and result in an increased level of DDG. Also, increased ethanol production means more corn is needed, which will increase corn prices.
9. The supply of DDG will eventually decrease feed prices since it is less costly than corn-based feed.
10. Corn prices affect feed prices. An increase in corn prices will increase feed prices, and *vice versa*.

of \$20 cwt, a midpoint of \$100 cwt, and an upper boundary of \$180 cwt. I expect participants to use this graphical framework when simulating outcomes for representation evaluation. Therefore, the higher \$165 (lower \$125) price is less (more) likely to incorporate the DDG-driven decline in the second half of the year. As a result, it is more likely that the \$165 (\$125) representation is inconsistent (consistent) when considering the entity-level evidence.

Control Variables

Individuals have varying abilities to process and manipulate information in working memory. To control for these variations, participants completed the short-form Raven Advanced Progressive Matrices Test, measuring their ability to manage sets of information in working memory (Raven 1976; Carpenter et al. 1990; Arthur and Day 1994). Additional control variables include self-assessments (using an 11-point scale) of mental effort on all packets, motivation on all packets, understanding of both the industry overview and the tutorial, and minutes spent on each packet.

Dependent Measures

Mental Model Development Measure

The predictions rely on the systems manipulation helping participants to develop more coherently organized mental models. To measure the success of this manipulation, I calculate mental simulation coherence, which captures the causal sequence structure of a mental model. I examine the extent that participants' mental simulation paragraphs have a logical causal chain throughout the paragraph. The paragraphs are coded based on the methodology used by Lesgold et al. (1988) to quantify experts' problem-solving methods because this measure captures participants' ability to perform a mental simulation that ties together components in a coherent progression without disconnections (Klein 1998). Participants list and then visually draw each attribute of the paragraph (i.e., each system component) on a graph. Exhibit 2 provides an example of the simulation paragraphs coding. The coherence metric is a percentage based on the number of system components connected by a single directional causal chain. This method ascribes a higher percentage to those simulations that tie together a larger number of components with a single directional causal chain.

Hypotheses Dependent Measures

To test H1, I evaluate the accuracy of participants' expectations of beef prices. The participants developed beef price expectations by sketching the time-series behavior of the prices for the period under review. I assign a point (maximum of 6 points) if the expectation sketches exhibited the same dynamic properties exhibited in a computer simulation (see Figure 2) using the facts and assumptions presented in the industry overview. Higher scores represent more consistency with the computer simulation.^{10,11}

If there are no significant differences between the industry-perspective conditions in regards to the amount of *necessary* information presented, then both conditions should have a basic understanding of the individual causal relations (i.e., the building blocks) driving beef prices. Any mental model improvements should be due to an inability to accurately organize these causal

¹⁰ I coded the simulation paragraphs with the assistance of another graduate student, covering all information regarding to which experimental condition the participants' packets were assigned. Both coders worked independently while coding, and mutually resolved any significant differences. The mean correlation index between raters for simulation coherence percentage is 0.85 (Kappa = 0.45, $p < 0.001$).

¹¹ A graduate student and I individually coded the graph scores and mutually resolved any significant differences. The mean correlation index between raters is 0.97 (Kappa = 0.81, $p < 0.001$).

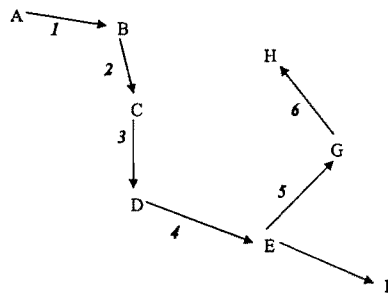
EXHIBIT 2

Example of Mental Simulation Scoring

The items below represent an example of how the mental simulation paragraphs were coded. Each component within the simulation was assigned a letter. Then causal linkages were established based on how the participant described the relationship between components. The coherence measure was calculated based on the number of identified components in the system and the longest directional chain through the system.

Number of Components: The total number of system components identified in the simulation (lettered items below).

Longest Chain: The longest causal chain that exists throughout the simulation (numbered items below).



- A. Increase in ethanol prices
 - B. Increase in corn prices
 - C. Increase in feed prices
 - D. Rise in beef prices
 - E. The company's cost of production
 - F. The company's use of substitute products
 - G. Decrease in beef consumption
 - H. The company's profits
- Number of Components: 8
 Longest Chain: 6
 Simulation Coherence (Longest Chain/# of Components): 0.75

relations, not from a lack of knowledge content of the causal relations. During the post-experimental questionnaire, participants answered multiple-choice questions concerning the relevant causal relations needed for developing accurate beef price expectations.

To test H2, I employ a commonly used workload efficiency measure that establishes a zero-efficiency benchmark and bases performance on a certain effort level (Clark et al. 2006). Participants provided a self-assessment rating of their mental effort level during the simulation paragraph task (using an 11-point scale). Such measures of mental effort are sensitive to variations in exerted effort (see Eggemeier and Wilson [1991] for review). I use simulation coherence as the performance measure in calculating processing efficiency because it represents the participants' level of attainment during their mental simulations (Paas et al. 2003).

To determine processing efficiency, I first recalculate the mental effort rating and task performance measures as z-scores. Then, I calculate processing efficiency as:¹²

$$EFF = \frac{P - R}{\sqrt{2}},$$

where *EFF* is the value of processing efficiency, *P* is the performance z-score, and *R* is the mental effort z-score. Both mental effort (x-axis) and performance (y-axis) are graphed on a Cartesian coordinate grid (see, e.g., Panel C of Table 3). *EFF* represents the perpendicular distance of a point on the axis $P - R$, which is the zero-efficiency line. In other words, on the line $P - R$, mental efficiency is zero ($E = 0$) because performance is at the expected level given the amount of effort exerted. For example, higher mental effort results in higher performance, and vice-versa. If $P - R > 0$, then *EFF* is positive, and if $P - R < 0$, then *EFF* is negative. Any point to the left of the zero-efficiency line ($E = 0$) represents an increase in processing efficiency, and any point to the right represents a decrease in processing efficiency. Experimental conditions are evaluated based on their relative position on the coordinate grid.

To test H3, I evaluate participants' credibility assessments (using an 11-point scale) of management representations regarding the company's decline in profitability. In addition, after evaluating management representations, the participants provided a quantitative range (i.e., confidence interval) for their beliefs regarding year-end beef prices. This measure accomplishes two objectives. First, by analyzing whether the participants' credibility assessments influence their ranges, I support H3 by showing that participants who recognize the inconsistencies incorporate this finding into their ranges rather than anchoring their assessment on management's representation. Second, by showing that the participants' initial ranges include their pre-developed expectations, I establish a benchmark to support my measures for expectation updating (H4), as described in the next section.

To test H4, I present participants with a paragraph including *new information* about the poultry industry. This new information is that poultry, unlike beef, cannot use the DDG byproduct for feed. Correct assimilation of this information into an existing mental model shows that the demand for beef should rise because its substitute product, poultry, will not experience the DDG-related reduction in feed costs and subsequent price decline that beef will. The industry overview (Packet A) states that beef supply cannot increase in the short term. Quantitative ranges regarding beef prices (see H3 dependent measures) should shift upward at year-end, reflecting an increased demand for beef because of higher prices for chicken, a beef substitute. I verified this movement by adding these facts into a computer simulation. Participants provided another quantitative range (confidence interval) for beef prices after considering this new information. I expect systems participants with a coherent mental model to mentally simulate how the new evidence affects the existing model and increase their beef price estimates by shifting both ends of their quantitative range upward. I use an indicator variable to signify whether participants shift their ranges accordingly.¹³

¹² This equation is based on the formula for calculating the distance from a particular point to the line $ax + by + c = 0$ in a coordinate system with coordinates (x, y) (Paas et al. 2003).

¹³ I do not examine the magnitude of the changes since there is no information in the case that suggests a specific dollar value change should occur. The dollar change is dependent on any assumptions made by the participant during the initial quantitative range and subsequent mental simulation during updating. Using dollar value of the confidence interval change as a dependent measure would slant the results in favor of those who made the largest dollar changes, whether correct or incorrect.

IV. RESULTS

Descriptive Statistics of Control Variables

The systems participants took more time to complete both the tutorial ($t_{55} = 5.45$, $p = 0.001$) and Packet A ($t_{55} = 2.87$, $p = 0.006$) than did reductionist participants.¹⁴ This outcome is not surprising, considering that the tutorial presented system participants with complex material that is not a part of most educational programs. Consistent with the increase in time, the systems participants' self-assessments of both effort level ($t_{59} = 2.97$, $p = 0.004$) and motivation ($t_{59} = 2.36$, $p = 0.022$) are higher than those of reductionist participants. This outcome is also intuitive because the systems-thinking material requires that participants learn less familiar and more complex concepts than the concepts learned by the reductionist participants. Both groups provide responses indicating that they understood the tutorial and overview, and there is no significant difference between conditions. There also are no significant differences for time and effort between industry perspective conditions for Packets B and C.

Systems View Manipulation Check

I expect that systems participants will generate mental simulations of entity-level evidence that are more coherently organized than those created by reductionist participants. As shown in Table 1, systems participants provide mental simulation paragraphs with a higher percentage of components held together by a single causal chain ($F_{1,57} = 5.39$, $p = 0.012$).¹⁵ This result suggests the systems intervention is effective in aiding how the systems participants organized the entity-level evidence in the case materials.¹⁶

Hypothesis 1

H1 predicts that systems participants will develop more accurate expectations of beef prices than will reductionist participants. As shown in Table 2, Panel A the systems participants graph scores are significantly higher than those provided by the reductionist participants ($t_{56} = 2.63$, $p = 0.005$).¹⁷ I collapse the individual expectations sketches for each condition in Panel C of Table 2, which shows that systems participants better understand the nonlinear behavior of beef prices and anticipate the decrease in prices after June. A regression analysis (untabulated) reveals that a systems perspective significantly predicts lower beef price estimates at year-end ($t_{56} = -4.50$, $p < 0.001$). These results support H1.

¹⁴ The systems (reductionist) participants took 13.53 (8.23) and 20.21 (15.23) minutes to complete the tutorial and packet A, respectively. Using time as a control variable does not qualitatively change the reported results in this study.

¹⁵ Further analysis of the simulation data (not tabulated) shows that self-assessed motivation increases the number of components listed ($t_{38} = 2.28$, $p = 0.013$), the number of causal links between components ($t_{38} = 2.03$, $p = 0.023$), and the length of the longest causal chain ($t_{38} = 1.80$, $p = 0.038$). However, motivation does *not* affect participants' ability to tie all of the relevant components together into a coherent simulation ($t_{38} = 0.58$, $p = 0.283$). This result supports the notion that constraints in working memory capacity are the underlying cause of the systems participants' improved simulations.

¹⁶ Because coherence is a ratio, multiple sets of numbers could provide the same ratio even though they are tying together a different number of components (e.g., 3/4 and 6/8 will both equal 0.75). Since simulations with fewer components could be tied together with relatively little cognitive effort, I ran an additional analysis that excluded any observations with less than seven components (approximately the 25th percentile). When these items are excluded, the systems participants' scores are still significantly higher ($F_{1,48} = 6.00$, $p = 0.009$).

¹⁷ Because effort level was significantly higher for the systems-based participants, the interaction between effort and industry analysis perspective is included in the model. I tested the graph scores for multicollinearity, because there is an expected level of dependency between the industry perspective condition and the self-assessments of effort level. I correct for the multicollinearity using the partial orthogonalization method. All remaining analyses with multicollinearity receive similar corrections. Due to the dependent nature of the variables, I also performed a path analysis, finding quantitatively similar results.

TABLE 1
Manipulation Check: Mental Simulation Coherence

Panel A: Mean (Standard Deviation) Simulation Coherence Scores
Industry Analysis Perspective

Dependent Measure	Reductionist	Systems
Simulation Coherence Scores	0.60 (0.18) n = 30	0.70 (0.18) n = 31
Components used to calculate Simulation Coherence		
Number of Components	8.95 (1.90)	8.18 (2.30)
Longest Chain	5.21 (1.76)	5.76 (2.31)

Panel B: ANCOVA Analysis for Mental Simulation Coherence Scores

Variable	F_{1,57}	p-value
Industry Analysis Perspective	5.39	0.012
Effort (Packet B)	0.07	0.396
Raven score	0.02	0.440

Indicator (0/1) coding was used for Industry Analysis Perspective in this table (systems-based perspective = 1; reductionist-based = 0).

All p-values are one-tailed. See Exhibit 2 for coding example.

Variable Definitions:

Simulation Coherence Scores = (Longest Chain/Number of Components);

Number of Components = number of individual components identified in the simulation paragraph; and

Longest Chain = longest causal chain that exists throughout the simulation paragraph.

To determine if both conditions understand the basic knowledge content necessary to estimate beef prices, during the post-experiment questionnaire, participants answered two multiple-choice questions essential to recognizing the nonlinear behavior of beef prices. The questions concern how the DDG byproduct affects feed prices (first question), and how feed prices affect beef prices (second question). Approximately 71 percent of the systems participants correctly answer the first question compared to approximately 76 percent of the reductionist participants ($t_{58} = -0.42$, $p > 0.500$). Approximately 97 percent of both conditions answer the second question correctly ($t_{58} = 0.05$, $p > 0.500$). These results imply that, when prompted, the reductionist participants possess knowledge content of the general causal relations driving beef prices but are unable to incorporate this knowledge properly into their expectations. Such a result is consistent with existing research, finding that decision-makers typically can identify and understand individual cues but typically cannot combine them properly when working memory is overwhelmed (Serman 1989; Diehl and Serman 1995; Einhorn 2000). Furthermore, this result lends support for my prediction that the systems view improvements are from knowledge organization differences, not entirely from differences in essential knowledge content between the conditions.

Hypothesis 2

In H2, I predict that participants with a systems perspective will exhibit more efficient working memory processing when performing a mental simulation than will reductionist participants.

TABLE 2

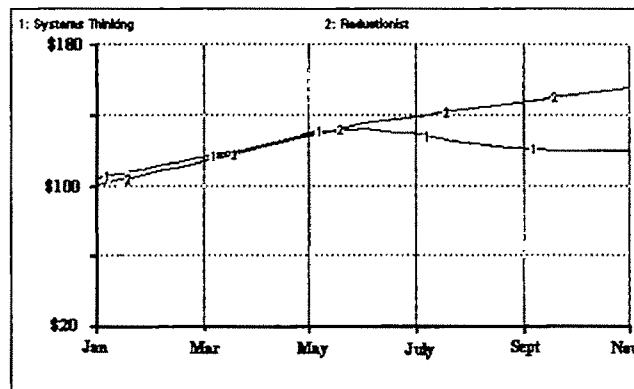
H1: Graph Scores for Beef Price Expectations

Panel A: Mean Graph Scores (Standard Deviation) of Beef Price Graph Scores
Industry Analysis Guidance

Dependent Measure	Reductionist	Systems
Graph Score (max of six)	2.35 (1.02) n = 30	3.25 (1.32) n = 31

Panel B: Regression Results for Graph Scores

Variable	β	SE	t	p-value
Constant	1.23	0.88	1.39	0.085
Industry Analysis Perspective	0.87	0.33	2.63	0.005
Effort (Packet A)	0.08	0.12	0.64	0.261
Effort * Industry Analysis Perspective	-0.08	0.24	-0.34	0.368
Raven score	0.08	0.06	1.20	0.117

Panel C: Graphical Representation of Expectations^a

Indicator (0/1) coding was used for Industry Analysis Perspective in this table (systems-based perspective = 1; reductionist-based = 0).

All p-values are one-tailed.

^a The graph was calculated by collapsing the participant's beef price graphs for each condition. The resulting average values were input into a systems software program to show the discrepancies between the expectations for each condition.

Variable Definitions:

Beef Price Graph Scores = how many points the participants' beef price expectations received when compared to a computer simulation of beef prices. See Figure 2 for the legend of the computer simulation of beef prices. A point was given if a participant's expectation graph exhibited the properties shown in Figure 2;

Effort = self-assessed mental effort level exerted in Packet A using an 11-point scale; and

Raven = total number of correct questions on the Raven's Progressive Matrices short-form test.

The participants' self-assessed understanding of both the client overview and the tutorial serve as control variables because an understanding of these items is essential for constructing an accurate simulation. Furthermore, because the Raven score controls for innate working memory processing ability, including this score in my analysis highlights any processing improvements stemming from mental model organization. The results in Table 3, Panel B show that the systems participants have more efficient processing when constructing a coherent mental simulation ($F_{1,55} = 4.60$, $p = 0.018$). Panel C provides a graphical presentation and shows that the systems-based efficiency score is above the zero-efficiency benchmark ($E = 0$) line, whereas the reductionist-based score is below the line. These results support H2 and are consistent with prior studies showing that increased use of causal schemas reduces working memory constraints and improves processing efficiency (Paas et al. 2003; Clark et al. 2006).

Hypothesis 3

H3 predicts that systems participants receiving inconsistent representations evaluate those representations as less credible than participants in the other three conditions. Table 4, Panel A provides descriptive statistics and shows a divergence of credibility assessments between the two industry analysis conditions. Panel B presents the conventional ANCOVA results controlling for effort level and Raven score. Because H3 predicts an ordinal interaction, contrast coding is the most appropriate method of testing the hypothesis (Buckless and Ravenscroft 1990). Contrast weights are +3 in the systems/inconsistent condition and -1 in the other three conditions. Panel C shows the planned contrast is significant since systems participants receiving inconsistent management representations provide the lowest credibility ratings ($F_{1,53} = 7.12$, $p = 0.005$). Follow-up tests of the simple effects show that systems auditors are able to distinguish between inconsistent and consistent representations ($F_{1,53} = 4.69$, $p = 0.017$), and are more likely than reductionist auditors to recognize inconsistent representations ($F_{1,53} = 6.62$, $p = 0.001$). Reductionist participants do not distinguish between evidence consistency, suggesting that they have difficulty simulating how the explanation compares to their existing knowledge. These results support H3.

H3 implies that systems participants will fixate their judgments less on management representations, and instead rely more by their own pre-developed expectations. I perform a regression analysis (untabulated) examining how the participants' credibility assessments affect their beef-price quantitative ranges. When the materials present inconsistent representations, the systems participants' credibility assessments are a significant predictor of the upper bound ($t_{48} = 2.26$, $p = 0.014$). This result suggests that those participants recognizing the inconsistency between their pre-developed expectations and management representations are less likely to anchor their inferences about beef prices on management representations (i.e., they are more likely to shift their confidence intervals downward in relation to the inconsistent management representation). By contrast, the reductionist participants' credibility assessments do not have a consistently significant effect on the parameters of the range. These results further support H3.

Hypothesis 4

H4 predicts that participants' most accurate updating of their mental models upon receiving new entity-level evidence will occur when they have a coherently organized mental model and when they take a systems perspective of the industry. Table 5, Panel A shows that the interaction of simulation coherence and systems-based condition significantly predicts whether participants

TABLE 3

H2: Processing Efficiency Results

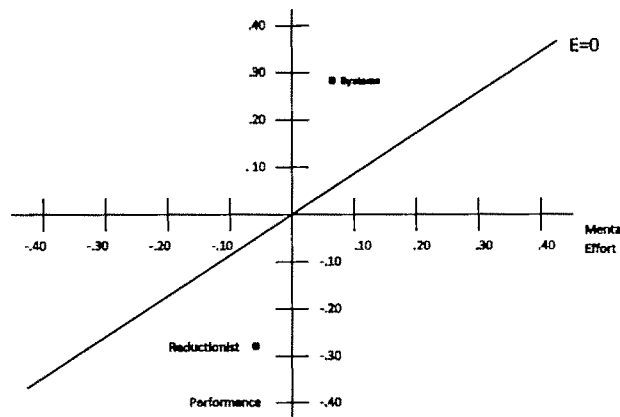
Panel A: Mean (Standard Deviation) Processing Efficiency Scores^a
Industry Analysis Guidance

Dependent Measure	Reductionist	Systems
Processing Efficiency Score	-0.17 (0.95) n = 30	0.17 (1.05) n = 31
Components of Processing Efficiency Score		
Performance Z-score	-0.29	0.29
Mental Effort Z-score	-0.50	0.50

Panel B: ANCOVA Analysis for Mental Simulation Coherence

Variable	F _{1,55}	p-value
Industry Analysis Perspective	4.60	0.018
Understanding—ST Tutorial	1.43	0.119
Understanding—BR Tutorial	1.15	0.144
Understanding—Overview	0.28	0.298
Raven	0.11	0.373

Panel C: Graphical Representation of Results



Indicator (0/1) coding was used for Industry Analysis Perspective in this table (systems-based perspective = 1; reductionist-based = 0).

All p-values are one-tailed.

^a The Processing Efficiency Score is the relative processing efficiency based on both performance and effort. In this calculation, the simulation coherence score (see Table 2) was used for the performance measures. Effort was the amount of self-assessed mental effort level exerted in Packet B.

Variable Definitions:

Processing Efficiency Score = [(Performance Z-score - Effort Z-score)/Square Root of 2];

Performance Z-score = calculated Z-score based on the simulation coherence score;

Mental Effort Z-score = calculated Z-score based on the self assessment of effort in Packet B;

Understanding—ST Tutorial = self-assessed understanding of the stock and flow tutorial given in the systems thinking condition using an 11-point scale;

(continued on next page)

Understanding—BR Tutorial = self-assessed understanding of the business risk tutorial given in the reductionist condition using an 11-point scale; and

Understanding—Overview = self-assessed understanding of the industry and company overview given to all participants using an 11-point scale.

increase both ends of the range (Wald $\chi^2_{1,52} = 3.12$, $p = 0.039$).¹⁸ I verify the robustness of these results by finding that the processing efficiency measure from H2 (see Panel B of Table 5) yields consistent results (Wald $\chi^2_{1,52} = 4.36$, $p = 0.019$). These results support H4.¹⁹

In sum, the results of this study show that a systems view improves information-processing ability and therefore the likelihood of acquiring and updating both knowledge content and structure. Most accounting studies assume that the effect of processing ability on knowledge acquisition is relatively static and does not change over the short-run (Libby 1995; Bonner 2007), but studies outside of accounting call for research that considers the implications of more dynamic associations between knowledge and ability (Ceci et al. 2003). My results suggest that a more dynamic relationship exists in the short-term between one component of knowledge, structure, and one component of information-processing ability, working memory capacity. This reinforcing relationship between knowledge organization and processing ability is significant because ignoring it fails to account for substantial variation in task performance (e.g., understanding and modeling complex business environments).

V. CONCLUDING REMARKS

This study examines whether a systems-based perspective of a client economic environment improves novice auditors' knowledge acquisition and subsequent analytical procedures performance. Auditing theorists suggest that training auditors to use systems-thinking skills can improve how they learn and use complex entity-level evidence that would normally overwhelm auditors (Bell et al. 1997; Bell et al. 2002). There is little empirical evidence showing the effectiveness of system approaches, and researchers and practitioners have struggled to find a successful and efficient way to implement systems skills due to the extensive training that is typically necessary.

I conduct a laboratory experiment in which novice participants learn about complex entity-level evidence for use in analytical procedures. I find that participants taking a systems view of the client environment develop more accurate and coherent mental models than participants using a reductionist perspective. In turn, systems participants are more likely to identify management

¹⁸ In Panel A, the Raven score significantly reduces the chances of increasing the confidence interval (Wald $\chi^2_{1,52} = 5.97$, $p = 0.007$), suggesting that participants with higher innate processing ability are more likely to *incorrectly* reduce the confidence interval. A further examination (untabulated) finds that the Raven score is significant for the reductionist-based condition (Wald $\chi^2_{1,52} = 3.58$, $p = 0.029$) but not for the systems-based condition (Wald $\chi^2_{1,52} = 1.22$; $p = 0.135$). To lend some insight into this result (untabulated), I find that the Raven score increases reductionist participants' self-assessed confidence level in evaluating beef prices ($t_{55} = 1.46$, $p = 0.075$) but decreases the systems participants' confidence level ($t_{55} = -1.44$, $p = 0.078$). This distinction implies that those with higher "fluid intelligence" levels, but with an impoverished mental model, are more likely to be overconfident in their ability to understand and utilize industry evidence. Conversely, those with better-organized knowledge structures recognize the complexity of the problem and are not overconfident in their ability to interpret it. This result supports the findings of Choy and King (2005), who suggest that decision-makers over-estimate their own innate abilities when solving dynamic problems.

¹⁹ I also examine other attributes of the updated beef price confidence intervals as dependent measures. The interaction of simulation coherence and systems thinking is either significant or marginally significant when using a separate indicator variable for an increase in the lower bound (Wald $\chi^2_{1,52} = 3.44$, $p = 0.032$) and the upper bound (Wald $\chi^2_{1,52} = 2.43$, $p = 0.059$) of the confidence interval. With these two measures, neither simulation coherence nor systems thinking are, in isolation, significant. Finally, I examine whether there was a tendency to *incorrectly reduce* the confidence intervals, finding no significant results for any condition.

TABLE 4
H3: Credibility Assessment of Management Representation

Panel A: Mean (Standard Deviation) of Credibility Assessments

	Industry Analysis Perspective		
	Reductionist	Systems	Totals
Consistent—\$125	7.07 (1.33) n = 14	7.57 (2.06) n = 14	7.32 (1.72) n = 28
Inconsistent—\$165	7.25 (1.49) n = 16	6.17 (1.53) n = 15	6.73 (1.67) n = 31
Totals	7.16 (1.49) n = 30	6.84 (1.91) n = 29	7.01 (1.70) n = 59

Panel B: Conventional ANCOVA Analysis

Variable	Sum of Squares	df	Mean Square	F	p-value
Industry Analysis Perspective (IAP)	2.307	1	2.307	0.89	0.175
Representation consistency (RC)	6.627	1	6.627	2.54	0.058
IAP * RC	11.203	1	11.203	4.30	0.022
Effort (Packet C)	13.959	1	13.959	5.36	0.012
Raven score	0.000	1	0.006	0.00	0.497
Error	138.031	53	2.604		

Panel C: Contrast Analysis and Follow-Up Simple Effects Test

Contrast	DF	F	p-value
Summary Hypothesis: Combination of inconsistency and systems-based perspective difference from other three conditions.	1,53	7.12	0.005
Follow-Up Simple Effect Tests			
Simple effect of Industry Analysis Perspective given inconsistent management explanation	1,53	6.62	0.001
Simple effect of consistent explanation versus inconsistent explanation given systems-based perspective	1,53	4.69	0.017

Indicator (0/1) coding was used for Industry Analysis Perspective in this table (systems-based = 1; reductionist-based = 0). Indicator (0/1) coding was also used for Representation Consistency (inconsistent = 1; consistent = 0).

All p-values are one-tailed equivalent.

Variable Definitions:

Effort = the self-assessed mental effort level exerted in Packet C using an 11-point scale;

Raven = the total number of correct questions on the Raven's Progressive Matrices short form test.

TABLE 5
H4: Updating of Mental Model
 Increase Confidence Interval = $\alpha + \beta_1(\text{Anchor}) + \beta_2(\text{Industry Analysis Perspective}) + \beta_3(\text{Simulation Coherence})$
 $+ \beta_4(\text{Simulation Coherence} * \text{Industry Analysis Perspective}) + \beta_5(\text{Raven}) + \beta_6(\text{Effort}) + \varepsilon$.

Panel A: Regression Results for Simulation Coherence			
Variable	β	SE	p-value
Constant	2.00	5.83	0.310
Anchor	-0.03	0.03	0.145
Industry Analysis Perspective (IAP)	-2.56	2.60	0.162
Simulation Coherence	4.07	5.33	0.223
Simulation Coherence * IAP	18.70	10.59	0.039
Raven	-0.67	0.27	0.007
Effort (Packet C)	0.31	0.44	0.236
Panel B: Regression Results for Processing Efficiency			
Variable	β	SE	p-value
Constant	4.62	5.14	0.185
Anchor	-0.02	0.03	0.235
Industry Analysis Perspective (IAP)	-1.11	1.82	0.272
Processing Efficiency	0.47	0.92	0.307
Processing Efficiency * IAP	3.62	1.74	0.019
Raven	-0.72	0.27	0.007
Effort (Packet C)	0.19	0.47	0.351

All p-values are one-tailed.

Variable Definitions:

Increase Confidence Interval = indicator variable that represents whether the participant increased both ends of their confidence intervals; coded 1 if both bounds are increased, 0 otherwise;

Processing Efficiency = processing efficiency score used to test H2;

Effort = self-assessed mental effort level exerted in Packet C using an 11-point scale; and

Raven = total number of correct questions on the Raven's Progressive Matrices short-form test;

Anchor = the beef price amount provided in management's explanation.

representations that are inconsistent with their expectations of entity-level evidence. In addition, the better-organized mental models of the systems participants provide more efficient use of working memory capacity, which aids the accurate updating of their expectations when in the presence of new industry-level evidence. These results are consistent with educational psychology findings that individuals can learn more effectively when they construct the incoming evidence as causal schematic structures (Clark et al. 2006). This study provides insight into how knowledge structure and processing ability interact to enhance learning, shortening the path to building auditors' expert-consistent performance when performing complex audit tasks (Nelson and Tan 2005).

Professional standards require auditors to understand complex entity-level evidence (AICPA 2006). Therefore, this study highlights a potential disconnect between what standards now require of auditors and what they are cognitively capable of doing. This disconnect suggests that new systems-based training methods could help auditors to learn better in the field and develop the skill sets needed to understand complex environments (PwC 2003; Peecher et al. 2007). Leaders in systems-thinking suggest that quantitative model building is the most effective way to fully learn and understand dynamically complex systems (Hannon and Ruth 1997; Sterman 2000). Future research could examine if new tools and decision aids, such as dynamic modeling, can help auditors accelerate learning about complex client environments.

Limitations

A potential concern with this study is whether my systems-based manipulation could be cost prohibitive in an actual audit because it would require the provision of pre-developed diagrams for each client. Recall that the purpose of this study is to examine whether and how systems techniques can improve audit quality, not to provide a specific methodology for firms. The systems manipulation is simply a tool to examine the improvements of a systems-consistent mental model. Therefore, while the diagram utilized in this study was necessary to help participants develop a systems view, I do not expect such a pre-developed diagram to be necessary in an actual audit. Systems advocates suggest the type of information provided in the systems manipulation could be either cognitively or formally built by someone properly trained in such skills (Hannon and Ruth 1997; Sterman 2000; Bell et al. 2002). Whether auditors can effectively and efficiently actually build these systems-consistent mental models remains a question for future research.

REFERENCES

- Ahn, W., W. F. Brewer, and R. J. Mooney. 1992. Schema acquisition from a single example. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 18 (2): 391–412.
- American Institute of Certified Public Accountants (AICPA). 2006. *Understanding the Entity and Its Environment and Assessing the Risks of Material Misstatement*. Statement on Auditing Standard No. 109. New York, NY: AICPA.
- Arthur, W., Jr., and D. V. Day. 1994. Development of a short form for the Raven Advanced Progressive Matrices Test. *Educational and Psychological Measurement* 54 (2): 394–403.
- Baddeley, A. 1998. *Human Memory: Theory and Practice*. Boston, MA: Allyn and Bacon.
- Bell, T. B., F. O. Marrs, I. Solomon, and H. Thomas. 1997. *Auditing Organizations Through A Strategic Systems Lens: The KPMG Business Measurement Process*. Montvale, NJ: KPMG LLP.
- , M. E. Peecher, and I. Solomon. 2002. *The Strategic-Systems Approach to Auditing: Cases in Strategic-Systems Auditing*. Montvale, NJ: KPMG LLP.
- , ———, and ———. 2005. *The 21st Century Public Company Audit: Conceptual Elements of KPMG's Global Audit Methodology*. Montvale, NJ: KPMG LLP.
- Bonner, S. E., and P. L. Walker. 1994. The effects of instruction and experience on the acquisition of auditing knowledge. *The Accounting Review* 69 (1): 157–178.

- , R. Libby, and M. W. Nelson. 1997. Audit category knowledge as a precondition to learning from experience. *Accounting, Organizations and Society* 22 (5): 387–410.
- . 2007. *Judgment and Decision Making in Accounting*. Upper Saddle River, NJ: Pearson Prentice Hall.
- Borthick, A. F., M. B. Curtis, and R. S. Siram. 2006. Accelerating the acquisition of knowledge structure to improve performance in internal control reviews. *Accounting, Organizations and Society* 31: 323–342.
- Brewer, W. F., and G. V. Nakamura. 1984. The nature and functions of schemas. In *Handbook of Social Cognition*, edited by R. S. Wyer, Jr., and T. K. Srull, Vol. 1: 119–160. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Buckless, F. A., and S. P. Ravenscroft. 1990. Contrast coding: A refinement of ANOVA in behavioral analysis. *The Accounting Review* 65 (4): 933–945.
- Carpenter, P. A., M. A. Just, and P. Shell. 1990. What intelligence test measures: A theoretical account of the processing in the Raven Progressive Matrices Test. *Psychological Review* 97 (3): 404–431.
- Ceci, S. J., S. M. Barnett, and T. Kanaya. 2003. Developing childhood proclivities into adult competencies: The overlooked multiplier effect. In *The Psychology of Abilities, Competencies, and Expertise*, edited by E. L. Grigorenko, and R. J. Sternberg, 70–91. London, U.K.: Cambridge University Press.
- Choo, F., and K. T. Trotman. 1991. The relationship between knowledge structure and judgments for experienced and inexperienced auditors. *The Accounting Review* 66 (3): 464–485.
- Choy, A. K., and R. R. King. 2005. An experimental investigation of approaches to audit decision making: An evaluation using systems-mediated mental models. *Contemporary Accounting Research* 22 (2): 311–350.
- Clark, R., F. Nguyen, and J. Sweller. 2006. *Efficiency in Learning: Evidence-Based Guidelines to Manage Cognitive Load*. San Francisco, CA: Pfeiffer.
- Conway, A. R. A., M. J. Kane, and R. W. Engle. 2003. Working memory capacity and its relation to general intelligence. *Trends in Cognitive Sciences* 7 (12): 547–552.
- Curtis, E., and S. Turley. 2007. The business risk audit—A longitudinal case study of an audit engagement. *Accounting, Organizations and Society* 32 (4–5): 439–461.
- Diehl, D., and J. D. Sterman. 1995. Effects of feedback complexity on dynamic decision making. *Organizational Behavior and Human Decision Processes* 62 (2): 198–215.
- Earley, C. E. 2001. Knowledge acquisition in auditing: Training novice auditors to recognize cue relationships in real estate valuation. *The Accounting Review* 76 (1): 81–97.
- . 2002. The differential use of information by experienced and novice auditors in the performance of ill-structured audit tasks. *Contemporary Accounting Research* 19 (4): 595–614.
- , V. B. Hoffman, and J. J. Joe. 2008. Reducing management's influence on auditors' judgments: An experimental investigation of SOX 404 assessments. *The Accounting Review* 83 (6): 1461–1485.
- Eggemeier, F. T., and G. F. Wilson. 1991. Performance-based and subjective assessment of workload in multi-task environments. In *Multiple-Task Performance*, edited by D. L. Damos, 217–278. Bristol, PA: Taylor & Francis Ltd.
- Einhorn, H. J. 2000. Expert judgment: Some necessary conditions and an example. In *Judgment and Decision Making: An Interdisciplinary Reader*, edited by T. Connolly, H. R. Arkes, and K. R. Hammond, 324–335. London, U.K.: Cambridge University Press.
- Engle, R. W., and M. J. Kane. 2003. Executive attention, working memory, capacity, and a two-factor theory of cognitive control. *Psychology of Learning and Motivation* 44: 145–199.
- Epley, N., and T. Gilovich. 2006. The anchoring and adjustment heuristic: Why the adjustments are insufficient. *Psychological Science* 17 (4): 311–318.
- Friedman, D., and S. Sunder. 1994. *Experimental Methods: A Primer for Economists*. New York, NY and London, U.K.: Cambridge University Press.
- Hannon, B., and M. Ruth. 1997. *Modeling Dynamic Biological Systems*. New York, NY: Springer.
- Hecht, G. 2010. The influence of auditor's understanding of circular causality on unintended consequence identification and control risk assessments. Working paper, Emory University.
- Hirst, D. E., and L. Koonce. 1996. Audit analytical procedures: A field investigation. *Contemporary Accounting Research* 13 (2): 457–486.
- Kahneman, D., and A. Tversky. 1982. The simulation heuristic. In *Judgment and Uncertainty: Heuristics and*

- Biases*, edited by D. Kahneman, P. Slovic, and A. Tversky, 201–208. London, U.K.: Cambridge University Press.
- Kennedy, J. 1995. Debiasing the curse of knowledge in audit judgment. *The Accounting Review* 70 (2): 249–272.
- Klein, G. 1998. *Sources of Power: How People Make Decisions*. Cambridge, MA: The MIT Press.
- Knechel, W. R. 2007. The business risk audit: Origins, obstacles, and opportunities. *Accounting, Organizations and Society* 32 (4-5): 383–408.
- Koonce, L. 1993. A cognitive characterization of audit analytical review. *Auditing: A Journal of Practice & Theory* 12 (Supplement): 57–76.
- Kyllonen, P., and D. Woltz. 1989. The role of cognitive factors in the acquisition of a cognitive skill. In *Abilities, Motivation, & Methodology: The Minnesota Symposium on Learning and Individual Differences*, edited by R. Kanfer, P. Ackerman, and R. Cudeck, 239–280. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Lesgold, A., R. Glaser, H. Rubinson, D. Klopfer, P. Feltovich, and Y. Wang. 1988. Expertise in a complex skill: Diagnosing x-ray pictures. In *The Nature of Expertise*, edited by M. T. H. Chi, R. Glaser, and M. J. Farr, 311–342. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Libby, R. 1995. The role of knowledge and memory in audit judgment. In *Judgment and Decision-Making Research in Accounting and Auditing*, edited by R. H. Ashton, and A. H. Ashton, 176–206. London, U.K.: Cambridge University Press.
- , R. Bloomfield, and M. W. Nelson. 2002. Experimental research in financial accounting. *Accounting, Organizations and Society* 27: 775–810.
- Moreno, K. K., S. Bhattacharjee, and D. M. Brandon. 2007. The effectiveness of alternative training techniques on analytical procedures performance. *Contemporary Accounting Research* 24 (3): 983–1014.
- Mussweiler, T., and F. Strack. 2001. The semantics of anchoring. *Organizational Behavior and Human Decision Processes* 86 (2): 234–255.
- Nelson, M., and H. T. Tan. 2005. Judgment and decision making research in auditing: A task, person, and interpersonal interaction perspective. *Auditing: A Journal of Practice & Theory* 24 (Supplement): 41–71.
- O'Donnell, E., and J. Schultz. 2005. The halo effect in business risk audits: Can strategic risk assessment bias auditor judgment about accounting details? *The Accounting Review* 80 (3): 921–938.
- Paas, F., J. E. Touvinen, H. Tabbers, and P. W. M. Van Gerven. 2003. Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist* 38 (1): 63–71.
- Peecher, M. E., and I. Solomon. 2001. Theory and experimentation in studies of audit judgments and decisions: Avoiding common research traps. *International Journal of Auditing* 5: 193–203.
- , R. Schwartz, and I. Solomon. 2007. It's all about audit quality: Perspectives on strategic-systems auditing. *Accounting, Organizations and Society* 32 (4-5): 463–485.
- PricewaterhouseCoopers (PwC). 2003. *Educating for the Public Trust: The PricewaterhouseCoopers Position on Accounting Education*. New York, NY: PricewaterhouseCoopers.
- Raven, J. C. 1976. *Advanced Progressive Matrices Set II*. San Antonio, TX: Harcourt Assessments, Inc. Adapted and reproduced with permission. All rights reserved. "Raven's Progressive Matrices" is a trademark of Harcourt Assessment, Inc. registered in the United States of America and and/or other jurisdictions.
- Reason, J. 1990. *Human Error*. London, U.K.: Cambridge University Press.
- Richmond, B. 1994. Systems thinking/system dynamics: Let's just get on with it. *System Dynamics Review* 10 (2-3): 135–157.
- Rickheit, G., and L. Sichelschmidt. 1999. Mental models: Some answers, some questions, some suggestions. In *Mental Models in Discourse Processing and Reasoning*, edited by G. Rickheit, and C. Habel, 9–40. Amsterdam, The Netherlands: Elsevier Science.
- Sanna, L. J. 2000. Mental simulation, affect, and personality: A conceptual framework. *Current Directions in Psychological Science* 9 (5): 168–173.
- Securities and Exchange Commission (SEC). 2003. *Report Pursuant to the Section 704 of the Sarbanes-Oxley Act of 2002*. Washington, D.C.: SEC.

- Smith, V. L. 1989. Theory, experiment and economics. *The Journal of Economic Perspectives* 3 (1): 151–169.
- Sterman, J. D. 1989. Misperceptions of feedback in dynamic decision making. *Organizational Behavior and Human Decision Processes* 43 (3): 301–335.
- . 2000. *Business Dynamics: Systems Thinking and Modeling for a Complex World*. Boston, MA: Irwin McGraw-Hill.
- Sternberg, R. 1997. Cognitive conceptions of expertise. In *Expertise in Context*, edited by P. Feltovich, K. Ford, and R. Hoffman, 149–162. Menlo Park, CA: The MIT Press.
- Sweeney, L. B., and J. D. Sterman. 2000. Bathtub dynamics: Initial results of a systems thinking inventory. *System Dynamics Review* 16: 249–286.
- Sweller, J., J. Van Merriënboer, and F. Paas. 1998. Cognitive architecture and instructional design. *Educational Psychology Review* 10 (3): 251–296.
- Tindall-Ford, S., P. Chandler, and J. Sweller. 1997. When two sensory models are better than one. *Journal of Experimental Psychology. Applied* 3: 257–287.
- Vera-Muñoz, S. C., M. Shackell, and M. Buehner. 2007. Accountants' usage of causal business models in the presence of benchmark data: A note. *Contemporary Accounting Research* 24 (3): 1015–1036.



Changes over Time in the Revenue-Expense Relation: Accounting or Economics?

Dain C. Donelson

Ross Jennings

John McInnis

The University of Texas at Austin

ABSTRACT: Dichev and Tang (2008) document a dramatic decrease over the last 40 years in the contemporaneous correlation between revenue and expense, along with an associated increase in earnings volatility and a decline in earnings persistence, suggesting a decline in earnings quality. We document that these changes are primarily attributable to an increase in the incidence of large special items. We then examine the extent to which this increase in special items is due to either more frequent real economic events related to special item recognition or to the adoption of new accounting standards. Our evidence suggests that changes in the frequency of economic events associated with special items have played a more important and sustained role relative to the role played by adoption of individual accounting standards. Finally, we find that the changing incidence of these economic events is at least in part related to the well-documented increase in competition in the U.S. economy over the last four decades.

Keywords: *earnings persistence; earnings volatility; matching; special items.*

Data Availability: *Data are available from public sources as identified in the text.*

JEL Classifications: *M41.*

I. INTRODUCTION

Dichev and Tang (2008; hereafter, DT) provide important evidence on changes over time in the relation between contemporaneous revenue and expense. For the 1,000 largest U.S. firms from 1967 to 2003, they find a decrease in the correlation between current-period revenue and current-period expense and an increase in the correlation between current-period revenue and expense from the prior and subsequent periods. These findings suggest a decline in “matching,” such that an increasing amount of expense is being recognized before and after the period in which it affects revenue. DT suggest these changes could arise from either changes in economic factors or changes in accounting standards, and provide indirect evidence that changes

We thank Bill Kinney, Laura Wang, Yong Yu, Mark Trombley (editor), two anonymous referees, and workshop participants at the Chinese University of Hong Kong and The University of Texas at Austin for their helpful comments.

Editor's note: Accepted by Mark Trombley.

*Submitted: August 2009
Accepted: November 2010
Published Online: April 2011*

in accounting standards are primarily responsible. In this study, we build upon DT by first identifying which expense line items are responsible for changes in the revenue-expense relation. We then examine whether the responsible line items have been more affected by changes in certain economic activity or by changes in specific accounting standards.

Understanding the factors leading to changes in the revenue-expense relation is important for researchers, financial statement users, and accounting standard-setters for two reasons. First, these changes have important implications for earnings quality. As modeled by DT and supported by their empirical results, reporting related revenue and expense in different accounting periods adds noise to the measurement of earnings. This noise increases earnings volatility and decreases earnings persistence, consistent with a decline in earnings quality (see, e.g., Dechow 1994; Francis et al. 2004; Dechow and Schrand 2004; DT). Thus, the results in DT suggest a prolonged decline in the quality of earnings.

Second, identifying the source of these changes in earnings properties is important because the two explanations we investigate have different policy implications. If accounting standards are responsible for changes in the revenue-expense relation, then the resulting decline in the “usefulness” of accounting information is in direct opposition to the objectives of the Financial Accounting Standards Board’s (FASB) conceptual framework. If this is the case, it would be important to understand which standards are most responsible and consider corrective action. Alternatively, if changes in economic activity are responsible, then standard-setters should consider this possibility when they evaluate current and future standards.

To determine which expense items are responsible for changes in the revenue-expense relation over time, we disaggregate total expenses into: cost of goods sold; selling, general, and administrative expense; depreciation; taxes; other income/expenses; and special items. Using a simple decomposition framework, we show that the decline in the relation between revenue and current expense is attributable primarily to a single income statement line item—special items. This line item consists mostly of asset impairments, restructuring charges, and gains/losses from asset sales. Not surprisingly, current-period special items have a much lower level of association with current-period revenue than other expenses. In addition, the “weight” of special items as a component of total expense has increased with the incidence of special items over time, decreasing the relation between current revenue and total current expense. We find that the most important changes in revenue and expense patterns are substantially reduced (by 90 percent in many cases) by either (1) excluding firm-years that report large special items (those greater than 1 percent of total assets), (2) using operating expense after depreciation (which excludes special items) to define total expenses, or (3) adding back special items to net income.

Next, we examine whether the increasing incidence of special items is due to changes in economic activity or changes in specific accounting standards. First, we provide evidence on whether the increasing incidence of special items is related to changes in the frequency of specific economic events associated with recognition of special items. We identify such economic events using five variables: negative employee growth; merger and acquisition activity; discontinued operations; negative revenue growth; and operating losses. These five variables represent real economic events whose frequencies are not likely to be influenced by accounting standards or by bookkeeping practices. Rather, they likely capture economic circumstances directly or indirectly associated with asset impairments, restructuring activities, and asset sales. We combine these variables into an index and show that this index both (1) increases throughout our study period, and (2) explains significant cross-sectional variation in the incidence of special items at the firm level. This finding provides strong evidence that the increase in special items is related to changes in the incidence of underlying economic events.

Next, we search for accounting standards adopted during our study period that are related to specific types of special items and find relatively few standards of this type. Among those that we

find, many appear to be either (1) reactive to changes in economic activity and accounting practice rather than leading such changes, or (2) too late in the study period to explain increases in special items prior to their adoption. In addition, we find that total special items increased in periods when no new accounting standards were adopted. Moreover, targeted tests of specific special items provide little evidence that adoption of specific accounting standards affected the frequency of these special items. Only for asset impairments do we find that implementation of a standard was associated with an increase in special items after controlling for a general time trend and our index of economic activity.

We also compare the effects of the specific accounting standards and our index of economic events on the increase in total special items. We find that changes in economic events play a more important role in the increasing incidence of special items over time relative to the role played by the adoption of specific accounting standards.

Finally, due to the apparently important role of economic events, we examine whether increases in our index of economic events are related to increasing competitive pressures in the economy documented by prior studies. We identify two measures of competitive pressure: industry concentration ratio and average probability of bankruptcy for the industry. We combine these measures into an index of competitive pressure and find that this index is (1) increasing over time throughout our study period, and (2) positively associated in the cross-section with our index of economic events related to special item recognition.

This study makes several important contributions. First, the results indicate that the deterioration in the contemporaneous relation between revenue and expense is due primarily to a dramatic increase in the frequency of reported special items, not to systematic issues across multiple line items in the income statement. Second, we provide strong evidence that an increase in the frequency of economic events associated with special item recognition plays an important role in the increasing frequency of special items. Third, the results also indicate that changes in specific accounting standards related to special items play, at best, a secondary role in the increasing frequency of special items. Thus, a primary implication of our results for standard-setters is that they should be aware of the relation between changes in the economic environment and the increasing incidence of special items over time.

II. BENCHMARK CHANGES IN THE REVENUE-EXPENSE RELATION

In this section, we provide benchmark evidence consistent with DT that the correlation between current-period revenue and current-period expense has declined over time. We begin by selecting a sample that is generally consistent with the sample in DT. First, we identify the 1,000 largest firms each year from 1967 to 2005 on the basis of total assets. Like DT, we select relatively large companies to confine our focus to changes in earnings properties among economically significant firms. This selection also helps maintain comparability throughout the sample period, as results are unaffected by the proliferation of smaller public firms later in the sample period. We require firms to have data available to calculate average earnings over successive, non-overlapping, two-year intervals ("two-year earnings") and the volatility of two-year earnings over the preceding ten years. We also require firms to have non-missing values for cost of goods sold, SG&A expense, income tax expense, and operating income after depreciation. We exclude financial firms, due to the difficulty in interpreting conventional income statement components for these firms.¹ Finally, we eliminate observations in the extreme 1 percent of the distribution each year of

¹ When we include financial firms in our sample, the results for the main analyses below are inferentially identical.

two-year earnings, two-year revenue and expense, two-year earnings volatility, current revenue, current expense, last-year expense, and next-year expense.² The sample has 32,645 firm-year observations between 1967 and 2005.

Next, we estimate the following cross-sectional regression by year, which is identical to the regression reported in DT (2008, Table 3):

$$Rev_{i,t} = \gamma_0 + \gamma_1 Exp_{i,t-1} + \gamma_2 Exp_{i,t} + \gamma_3 Exp_{i,t+1} + \varepsilon_{i,t}. \quad (1)$$

$Rev_{i,t}$ is total revenue for firm i for year t , and $Exp_{i,t}$ is total expense, computed as revenue minus net income before extraordinary items for firm i in year t , divided by average total assets for year t . These, and all other variables used in the study, are defined in the Appendix. The coefficient on current-period expense estimates the extent to which current-period expense varies with revenue in the current period, and the coefficient estimates on expense in the prior period and the next period estimate the extent to which expenses are scattered to adjacent periods.

Average coefficients are reported in Table 1, Panel A, for the first half (1967 to 1985) and second half (1986 to 2005) of the study period. The average coefficient estimate on contempora-

TABLE 1
Properties of Earnings-Related Variables over Time

Panel A: Relation of Revenue to Lagged, Current and Future Expense

	<i>Exp_{t-1}</i>	<i>Exp_t</i>	<i>Exp_{t+1}</i>
1967–1985	0.002	1.032	–0.030
1986–2005	0.089	0.895	0.025
Difference	0.087***	–0.137***	0.055***
t-statistic	(5.98)	(–5.54)	(4.04)

Panel B: Volatility and Persistence of Earnings

	Volatility	Persistence
1967–1985	0.022	0.805
1986–2005	0.036	0.680
Difference	0.014***	–0.125***
t-statistic	(9.77)	(–6.32)

*, **, *** Represent significance at the 10 percent, 5 percent, and 1 percent levels, respectively, in two-tailed tests. For difference calculations only, t-statistics are in parentheses below the coefficient estimates.

This table presents properties of earnings-related variables between two time periods, 1967–1985 and 1986–2005. Panel A presents average annual coefficients from estimating the following model each in both time periods:

$$\text{Model 1: } Rev_{i,t} = \gamma_0 + \gamma_1 Exp_{i,t-1} + \gamma_2 Exp_{i,t} + \gamma_3 Exp_{i,t+1} + \varepsilon_{i,t}.$$

Panel B presents average firm specific volatility of *Earn* and average coefficients (persistence) from estimating the following model each year in both time periods:

$$\text{Model 2: } Earn_{i,t} = \gamma_0 + \gamma_1 Earn_{i,t-1} + \varepsilon_{i,t}.$$

Variables are defined in the Appendix.

² These eliminations follow the sample selection procedures in DT. When we do not eliminate these observations, the decrease in the contemporaneous correlation between revenue and expense is somewhat more pronounced.

neous expense is 1.032 (0.895) in the early (late) period, and the difference of 0.137 is highly significant ($p < 0.01$). We also report increases in the lag (lead) coefficient from 0.002 (−0.030) in the early period to 0.089 (0.025) in the late period.

As described by DT, the decline in the contemporaneous association of revenue and expense is likely to have implications for the volatility and persistence of earnings, which in turn have important implications for the quality of earnings. Specifically, scattering expenses related to current revenue to other time periods adds “noise” to reported expenses, resulting in more volatile and less persistent earnings. To see this, we follow DT and compute the firm-specific volatility of earnings (as measured by its standard deviation) over rolling five-year windows, and then average this measure across firms each year and then across years within each sub-period. We also follow DT and estimate earnings persistence using the following regression:

$$Earn_{i,t} = \gamma_0 + \gamma_1 Earn_{i,t-1} + \varepsilon_{i,t}. \quad (2)$$

The results for both earnings volatility and earnings persistence are presented in Panel B of Table 1.³ The volatility of earnings increases from 0.022 in the early period to 0.036 in the later period, and the change is significant ($p < 0.01$). The earnings persistence coefficient (γ_1 from model 2) decreases from 0.805 in the early period to 0.680 in the later period, and this difference is also significant ($p < 0.01$). These results indicate that the decline in the contemporaneous relation between revenue and expense documented in Panel A of Table 1 has produced an increase in the volatility of earnings and a decrease in the persistence of earnings from the early years to the later years in our study period, consistent with a decline in earnings quality (see, e.g., Dechow 1994; Francis et al. 2004; Dechow and Schrand 2004).

Overall, the results reported in Table 1 are virtually identical to those reported by DT (2008, Tables 3–5). In the next section we examine which income statement line items are responsible for the decline in the revenue-expense relation documented in Table 1.

III. DISAGGREGATED EXPENSES AND THE REVENUE-EXPENSE RELATION

We begin our analysis with a decomposition of the change in average coefficients from Equation (1) between the early period and the late period. To do this, we divide total expense in Equation (1) into six components: cost of goods sold (*COGS*), selling, general, and administrative expense (*SGA*), depreciation expense (*DEPR*), tax expense (*TAX*), other expenses (*OTH*), and special items (*SI*). *SGA* includes R&D and advertising expense. Estimating Equation (1) implicitly constrains all of the coefficients on these six components that comprise total expense for a given year to be identical. Alternatively, the following regression allows the coefficients on these six components for the current year to vary freely:

$$Rev_{i,t} = \gamma_0 + \gamma_1 Exp_{i,t-1} + \gamma_{2a} COGS_{i,t} + \gamma_{2b} SGA_{i,t} + \gamma_{2c} DEPR_{i,t} + \gamma_{2d} TAX_{i,t} + \gamma_{2e} OTH_{i,t} + \gamma_{2f} SI_{i,t} + \gamma_3 Exp_{i,t+1} + \varepsilon_{i,t} \quad (3)$$

A decomposition similar to that in Kee (2009) demonstrates that the annual coefficient estimate for Exp_t (γ_2) in Equation (1) is a weighted average of the parameters on the individual component estimates (γ_{2a} through γ_{2f}) from Equation (3).⁴ Thus, the decline we observe in the coefficient estimate on aggregate expense may occur for one or both of two main reasons. First, it

³ To mitigate the influence of extreme observations in our volatility and persistence tests, we eliminate observations in the extreme 1 percent of the distribution each year of earnings volatility and current- and prior-year earnings.

⁴ The decomposition in Kee (2009) is limited to two components that are assumed to be positively correlated. Our analysis is based on six components whose pair-wise correlations may be either positive or negative. A proof that the analysis in Kee (2009) extends to our case without loss of generality is available from the authors on request.

is possible that the coefficient estimates for some (or all) of the individual expense components decline between the two periods. Second, it is also possible that the “weights” of the individual components have shifted over time, away from components with larger coefficient estimates and toward components with smaller coefficient estimates. As demonstrated by Kee (2009), the weights are related to the importance of a particular component—specifically, how much of the variance of total expense is attributable to each expense component—and must sum to 1 by construction.⁵

Table 2 summarizes the changing coefficients and weights of the expense components. The first set of rows provides the average coefficient estimates for the expense components for the early and late periods as well as the change between periods. The next set of rows provides the average weight for each component for each period as well as the change between periods. It is clear from Table 2 that special items has the lowest average coefficient estimate in both periods and that its average coefficient estimate has declined the most from the early period (0.483) to the late period (0.197). At the same time, we observe a dramatic shift in weight toward special items (from 0.009 to 0.129) and away from high-coefficient components like *COGS* and *TAX*. The increased weight on special items is consistent with prior studies that document an increase over time in one-time, special charges such as asset write-offs, asset sales, and restructuring charges (Elliott and Hanna 1996; Collins et al. 1997; Fairfield et al. 2009).⁶

TABLE 2
Decomposition Analysis
Current Year Expense

		<i>COGS</i>	<i>SGA</i>	<i>DEPR</i>	<i>TAX</i>	<i>SI</i>	<i>OTH</i>
Coefficients	1967–1985	1.004	1.017	1.058	1.891	0.483	0.584
	1986–2005	0.996	1.030	0.853	2.123	0.197	0.480
	Difference	−0.009	0.014	−0.205	0.233	−0.286	−0.105
Weights	1967–1985	0.834	0.106	0.006	0.039	0.009	0.007
	1986–2005	0.711	0.122	0.013	0.011	0.129	0.013
	Difference	−0.122	0.016	0.008	−0.027	0.120	0.006

We decompose the aggregate coefficient estimates on current expense from Panel A of Table 1 into six income statement components and their respective weights, using the following model:

$$\text{Model 3: } Rev_{i,t} = \gamma_0 + \gamma_1 Exp_{i,t-1} + \gamma_{2a} COGS_{i,t} + \gamma_{2b} SGA_{i,t} + \gamma_{2c} DEPR_{i,t} + \gamma_{2d} TAX_{i,t} + \gamma_{2e} OTH_{i,t} + \gamma_{2f} SI_{i,t} + \gamma_3 Exp_{i,t+1} + \varepsilon_{i,t}$$

Kee (2009) demonstrates that the coefficient on current expense in Panel A of Table 1 (γ_2) is simply a weighted average of coefficients on the expense components (γ_{2a} through γ_{2f}) in (3). The weight for any component, equals $\text{Cov}(Exp, component_i) / \text{Var}(Exp)$, where Exp is total current expense (Kee 2009).

⁵ The weight for any component, equals $\text{Cov}(Exp, component_i) / \text{Var}(Exp)$, where Exp is total current expense (see Kee [2009] for a proof). This is the coefficient one obtains from an OLS regression of $component_i$ on total expense. Since $\text{Var}(Exp) = \sum_i \text{Cov}(Exp, component_i)$, the weights must sum to 1.

⁶ A recent study by Cready et al. (2010) documents that the persistence of special items varies positively with the frequency of reported special items in the prior three years, indicating that for some firms these nonrecurring items are more recurring than for other firms. However, their evidence indicates that even for firms with five or more special items

These results are consistent with the decline we observe for the coefficient estimate on contemporaneous expense in Equation (1)—the low (and decreasing) coefficient estimate on special items increases in importance (weight) and drags down the coefficient on total expense. In contrast, the results for the other expense components are not consistent with the decline in the coefficient estimate on contemporaneous expense in Equation (1). For the other expense components, either the change in the estimated coefficient from the early period to the late period is very small or increasing (*COGS* and *SGA*) or the weight of the component is low in both periods (*DEPR*, *TAX*, and *OTH*).⁷

We perform four analyses to corroborate the critical role of special items in changes in the revenue-expense relation and the volatility and persistence of earnings reported in Table 1. First, we repeat all analyses in Table 1 after deleting observations with a large special item that would affect any of the results, where “large” is defined as 1 percent of average total assets for the year.⁸ The results, which are reported in Table 3, Panel A, demonstrate the substantial effect large special items have on the earnings properties. After excluding observations with large special items, the coefficient estimate for current-period expense from Equation (1) is not significantly different between the early and late periods ($p = 0.78$), as compared to a significant decline of 0.137 in Table 1 ($p < 0.01$). In addition, between the early and late periods volatility (persistence) increases (decreases) by only 0.002 (0.003), as compared to 0.014 (0.125) in Table 1. Although the increase in volatility remains significant at the 0.01 level after excluding observations with large special items, the magnitude is only 14 percent of the magnitude reported in Table 1.

As a second method of evaluating the effect of large special items, we sort observations in the later period into those with and without large special items. There are not enough observations with large special items during the first years of the early period to conduct this analysis in the early period. The results, which are reported in Table 3, Panel B, demonstrate the dramatic effect of special items. For observations without special items, the late period coefficient estimate for contemporaneous expense from Equation (1) is 1.046, which is similar to the early period coefficient estimate reported in Table 1 (1.032). In contrast, the late period contemporaneous expense coefficient for observations with large special items is 0.752, which is significantly less than the estimate for observations without large special items ($t = 8.44$; $p < 0.01$). In addition, observations with large special items have significantly larger (smaller) estimates for volatility (persistence), demonstrating the effect of special items on these earnings properties.

in the previous twelve quarters, the persistence of those special items is still substantially below that of earnings adjusted for special items (see Cready et al. 2010, Table 3). Thus, while the recurring nature of some special items might partially mitigate the effects we report in Table 2, such mitigation is likely to be slight.

⁷ Although the weight on *COGS* declines substantially from the early to the late period in Table 2, we view the decline in the revenue-expense relation as a function more of special items rather than *COGS*. The weights must sum to 1, such that a dramatic increase in relative importance of special items must mechanically be offset by a decline in relative importance of other components. This issue is analogous to the product mix problem in management accounting where a company has a low-margin product (LM) and a high-margin product (HM). Assume that HM volume grows 10 percent and LM volume doubles, shifting the mix away from HM and toward LM and reducing overall profit margin. In this case it is logical to attribute the decline in profit margin to LM, the product whose growth was extreme. Similarly, in our case, we attribute the weight effect to special items, the component whose growth was extreme. In an untabulated analysis we find that more than 90 percent of the change in the coefficient estimate on aggregate expense in Equation (1) is due to the combined effects of the change in the coefficient estimate and the change in the weight of special items (details available from the authors on request).

⁸ Specifically, we re-estimate Equation (1) after deleting all observations with a large special item in the current year (5,330 observations), we re-estimate the volatility measures after deleting observations with a large special item in any of the prior five years (12,145 observations), and we re-estimate the persistence regression after deleting observations with large special items in the current or previous years (7,728 observations). In addition, we repeat these analyses with a smaller, common sample formed by the intersection of these three subsamples, and results are very similar. Our definition of “large” is consistent with that used by Elliott and Hanna (1996).

TABLE 3
Properties of Earnings-Related Variables over Time
Adjusting for Large Special Items

Panel A: Early versus Late Period, Excluding Observations with Large Special Items

	<i>Exp_{t-1}</i>	<i>Exp_t</i>	<i>Exp_{t+1}</i>
1967–1985 (No Large <i>SI</i>)	–0.003	1.042	–0.036
1986–2005 (No Large <i>SI</i>)	0.010	1.046	–0.046
Difference	0.013*	0.004	–0.010*
t-statistic	(1.97)	(0.42)	(–1.73)
	Volatility	Persistence	
1967–1985 (No Large <i>SI</i>)	0.020	0.822	
1986–2005 (No Large <i>SI</i>)	0.022	0.819	
Difference	0.002***	–0.003	
t-statistic	(3.11)	(–0.15)	

Panel B: Late Period, With versus Without Large Special Items

	<i>Exp_{t-1}</i>	<i>Exp_t</i>	<i>Exp_{t+1}</i>
1986–2005 (No Large <i>SI</i>)	0.010	1.046	–0.046
1986–2005 (Large <i>SI</i>)	0.165	0.752	0.076
Difference	0.155***	–0.294***	0.123***
t-statistic	(8.10)	(–8.44)	(5.53)
	Volatility	Persistence	
1986–2005 (No Large <i>SI</i>)	0.022	0.819	
1986–2005 (Large <i>SI</i>)	0.032	0.529	
Difference	0.010***	–0.290***	
t-statistic	(15.94)	(–11.43)	

Panel C: Early versus Late Period, Based on Operating Income

	<i>Exp_{t-1}</i>	<i>Exp_t</i>	<i>Exp_{t+1}</i>
1967–1985	–0.012	1.067	–0.046
1986–2005	0.011	1.059	–0.058
Difference	0.024**	–0.008	–0.012
t-statistic	(2.85)	(–0.63)	(–1.35)
	Volatility	Persistence	
1967–1985	0.035	0.819	
1986–2005	0.034	0.813	
Difference	–0.001	–0.006	
t-statistic	(–0.50)	(–0.15)	

(continued on next page)

Panel D: Early versus Late, Adding Back Special Items

	<u>Exp_{t-1}</u>	<u>Exp_t</u>	<u>Exp_{t+1}</u>
1967–1985	−0.003	1.038	−0.031
1986–2005	0.023	1.024	−0.038
Difference	0.025***	−0.013*	−0.007
t-statistic	(3.99)	(−1.70)	(−1.09)
	<u>Volatility</u>	<u>Persistence</u>	
1967–1985	0.021	0.822	
1986–2005	0.027	0.786	
Difference	0.006***	−0.036*	
t-statistic	(6.56)	(−1.87)	

*, **, *** Represent significance at the 10 percent, 5 percent, and 1 percent levels, respectively, in two-tailed tests.

For difference calculations only, t-statistics are in parentheses below the coefficient estimates.

This table presents properties of earnings-related variables for two time periods, 1967–1985 and 1986–2005, utilizing four different methods to control for the presence of special items. All regression specifications are equivalent to those estimated in Table 1 except as noted below.

In each panel, the first set of results presents average annual coefficients from estimating:

$$\text{Model 1: } Rev_{i,t} = \gamma_0 + \gamma_1 Exp_{i,t-1} + \gamma_2 Exp_{i,t} + \gamma_3 Exp_{i,t+1} + \varepsilon_{i,t}.$$

In each panel, the second set of results presents average firm-specific volatility of *Earn* and average coefficients (persistence) from estimating:

$$\text{Model 2: } Earn_{i,t} = \gamma_0 + \gamma_1 Earn_{i,t-1} + \varepsilon_{i,t}.$$

Panel A presents results for the early and late time periods after excluding firm-years with large special items (defined as greater than 1 percent of average total assets).

Panel B presents results for the late time period, comparing observations with large special items to observations without large special items.

Panel C presents results for the early and late time periods based on operating income after depreciation as the earnings measure.

Panel D presents results for the early and late time periods after adding back special items to earnings before discontinued operations.

Variables are defined in the Appendix.

Our third and fourth analyses avoid the deletion of observations and use alternative earnings measures to evaluate the effect of special items. In the third analysis, we replace net income before extraordinary items with operating income after depreciation (but before special items) and we replace total expense with revenue minus operating income after depreciation. In the fourth analysis, we add special items back to net income before extraordinary items and replace total expense with revenue minus this modified income figure.

The results for the third (fourth) analysis are reported in Panel C (D) of Table 3. In Panel C (D), the change in the coefficient estimate for current-period expense from Equation (1) from the early to the late period is −0.008 (−0.013), as compared to −0.137 in Table 1. The difference in Panel C is 6 percent of the original difference and is not significantly different from 0 ($p = 0.52$), while the difference in Panel D is 10 percent of the original difference and is weakly significant ($p = 0.10$). Also, in Panel C, the changes in both volatility and persistence from the early to the late period are small when compared to the results in Table 1, and neither change is statistically significant ($p = 0.61$ and 0.59 , respectively). In Panel D, when special items are added back to

earnings, the increase (decrease) in volatility (persistence) is 43 percent (29 percent) of the increase (decrease) reported in Table 1, and both changes are significant ($p < 0.01$; $p = 0.07$).⁹

We also observe that adjusting for special items in the current period affects the coefficient estimates on aggregate expense in the surrounding periods. Specifically, the increases in the coefficient estimates on non-contemporaneous expenses are attenuated or reversed. In Table 1, these coefficient estimates likely increased because the increased incidence of special items added noise to the contemporaneous relation between revenue and total current expense over time. This noise was likely compensated for by an increase in the relation between current revenue and non-current expense because (1) these expenses are highly correlated with prior period and next period revenue, and (2) revenue is highly persistent from one period to the next. Thus, all of the patterns we observed in Table 1 are substantially reduced by each of our adjustments for special items in Table 3.

Taken together, these results suggest that the significant patterns documented by DT and reported in Table 1 are due primarily to the increase in the weight of special items as a component of total expense. In the next sections, we provide further evidence on the increase in special items over time and explore explanations for this phenomenon.

IV. CHANGES OVER TIME IN THE INCIDENCE OF SPECIAL ITEMS

To provide a general view of the extent to which special items are increasing during the sample period, we first plot the percent of sample observations that reported a large special item for each year from 1967 to 2005 in Figure 1. The incidence of large special items is below 5 percent before 1973 and rises to almost 50 percent in 2001. This pattern is consistent with Elliott and Hanna (1996, Figure 1) and Fairfield et al. (2009, Figure 1).

Next, we provide additional descriptive evidence on changes in the frequency of commonly reported special items using data reported by *Accounting Trends and Techniques*, an annual publication of the American Institute of Certified Public Accountants. Each volume contains examples of corporate disclosures and tabulations of reporting practices for 600 industrial, retail, and service companies chosen from the Fortune 1000.¹⁰ Throughout most of our study period, *Accounting Trends and Techniques* has included a tabulation of various types of other revenue and gains and other expenses and losses disclosed by firms. We summarize these tabulations in Table 4, Panel A for gains and in Panel B for losses. We tabulate only items that the Compustat manual includes in its definition of special items. Each cell reports the percent of the 600 firms surveyed by *Accounting Trends and Techniques* reporting that item that year. The columns sum to more than 100 percent because a company may have more than one type of gain or loss (e.g., sale of assets and litigation settlement), and a gain or loss may count in more than one category (e.g., write-down of assets and restructuring of operations).¹¹

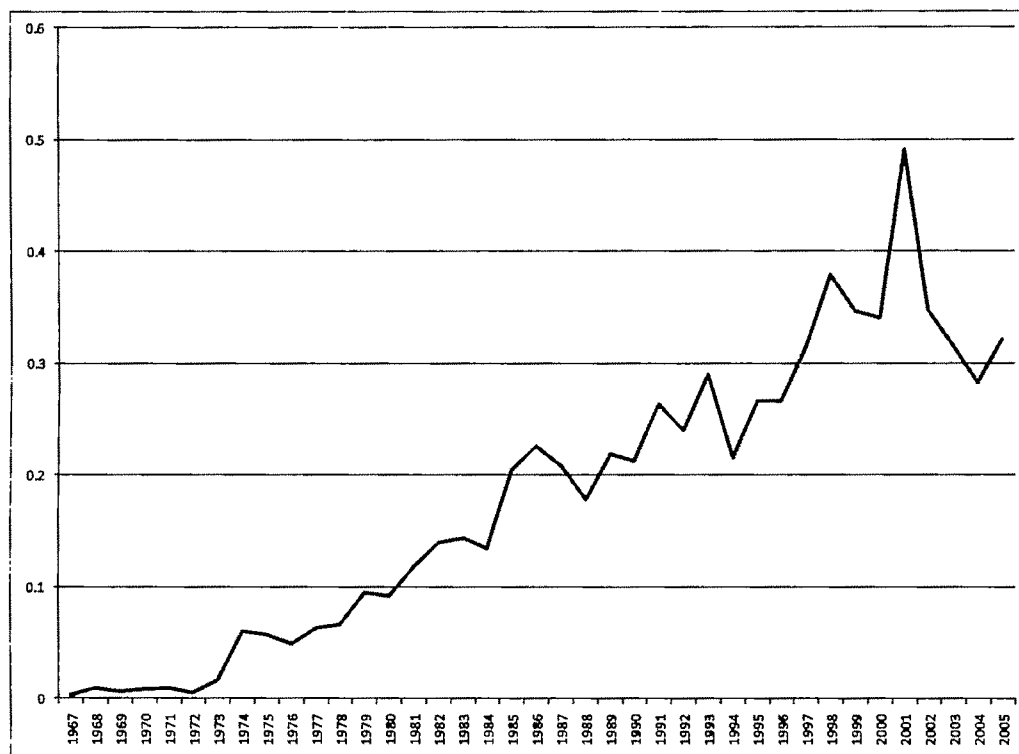
The results in Table 4, Panel A, indicate a steady increase in gain special items between 1975 and 2005, with average total incidence for the first (last) three years of 9 percent (54 percent),

⁹ The tests reported in Panel D of Table 3 are similar to a robustness test conducted by DT. DT report that the original change in the coefficient estimate on contemporaneous expense from Equation (1) of -0.149 falls to -0.081 when special items are added back to income. We find that our original change of -0.137 falls to -0.013 after adding back special items, which is consistent with our decomposition results in Table 2 and our results from the additional adjustments for special items in Table 3.

¹⁰ There is a substantial overlap between this group and our sample. For example, for 2002, of the 750 firms in our sample, 350 (47 percent) were also included in the firms surveyed by *Accounting Trends and Techniques*.

¹¹ Compustat and *Accounting Trends and Techniques* use similar procedures to identify special items. In both cases, items are identified as a special item when they are disclosed either in the main financial statements or in the financial statement notes. See Burgstahler et al. (2002) for discussion of Compustat procedures.

FIGURE 1
Large Special Items over Time



This figure graphs the incidence of large special items (those greater than 1 percent of average total assets) per year in the sample.

which is about a sixfold increase. Among the individual gain special items, only asset sales have a consistently high frequency, rising from 8 percent in 1975 to 37 percent in 2005.

The incidence of loss special items (Panel B) is rising over time more rapidly than the incidence of gain special items. The most common loss special items are asset write-downs, restructuring charges, and losses on asset sales. Across all loss special items, the average total incidence for the first (last) three years reported in the table is 5 percent (126 percent), which is more than a 25-fold increase. Even if we eliminate write-downs and asset sales as being included in restructurings in later years, loss special items still increase 16-fold. Loss special items are increasing faster than gain special items both because the types of items being reported are increasing and because of increases in the frequency of particular items being reported. For example, only two types of loss special items are reported in *Accounting Trends and Techniques* for 1980, while for 2005 eight types of loss special items have reported amounts. Also, in 1980 only 3 percent of firms report an asset write-down, but this percentage increases to 35 percent in 2005.

TABLE 4
Losses and Gains from Accounting Trends and Techniques

Panel A: Gains		05	04	03	02	01	00	99	98	97	96	95	94	93	92	91	90
Special Items (%)																	
Asset sale	37	33	33	33	32	30	16	26	25	27	21	23	20	17	20	18	22
	12	12	11	11	11	8	6	5	4	2	2	3	nc	nc	nc	nc	nc
	5	6	6	3	3	2	3	1	3	2	1	2	2	3	3	2	2
	1	*	nc	nc	nc	nc	nc	nc	nc	nc	nc	nc	*	1	1	2	5
	1	2	2	1	1	nc	nc	nc	nc	nc	nc	nc	nc	nc	nc	nc	1
Total	56	53	52	46	40	25	32	32	32	31	24	28	22	21	24	22	29
Panel A: Gains, continued		89	88	87	86	85	84	83	82	81	80	79	78	77	76	75	
Special Items (%)																	
Asset sale	27	24	25	26	25	25	20	22	20	20	16	15	12	12	11	8	8
	nc	nc	nc	nc	nc	nc	nc	nc	nc	nc	nc	nc	nc	nc	nc	nc	nc
	2	2	2	1	2	2	2	1	nc	nc	nc	nc	nc	nc	nc	nc	nc
	3	4	5	5	1	1	1	nc	nc	nc	nc	nc	nc	nc	nc	nc	nc
	1	1	1	1	3	5	5	4	10	nc	nc	nc	nc	nc	nc	nc	nc
Total	33	31	33	33	31	31	28	27	30	20	16	15	12	12	11	8	8
Panel B: Losses		05	04	03	02	01	00	99	98	97	96	95	94	93	92	91	90
Special Items (%)																	
Write-down	35	32	34	36	39	39	27	25	29	18	19	19	11	11	8	11	12
	12	12	13	16	12	12	7	8	5	7	5	4	4	5	6	7	6
	42	38	37	44	48	48	26	27	33	24	23	22	17	32	23	24	19
	10	10	8	7	6	6	4	4	6	4	6	4	5	4	3	3	3
	6	6	4	4	3	3	2	3	5	5	4	5	5	5	4	5	5
Purch R&D	3	3	2	3	4	4	3	3	7	3	nc	nc	nc	nc	nc	nc	nc
	12	10	12	nc	nc	nc	nc	nc	nc	nc	nc	nc	nc	nc	nc	nc	nc

(continued on next page)

Panel B: Losses

Special Items (%)	05	04	03	02	01	00	99	98	97	96	95	94	93	92	91	90
Debt ext	13	14	10	nc	nc	nc	nc	nc	nc	nc	nc	nc	nc	nc	nc	nc
Total	133	125	120	110	112	69	70	85	61	57	54	42	57	44	58	45

Panel B: Losses, continued

Special Items (%)	89	88	87	86	85	84	83	82	81	80	79	78	77	76	75
Write-down	8	7	2	10	8	7	6	6	5	3	3	2	3	3	0
Asset sale	5	4	5	5	5	4	5	4	3	4	4	5	5	4	0
Restructuring	16	12	14	22	17	12	13	15	9	8	nc	nc	nc	nc	nc
Litigation	2	3	2	2	3	2	2	nc	nc	nc	nc	nc	nc	nc	nc
Cleanup	nc	nc	nc	nc	nc	nc	nc	nc	nc	nc	nc	nc	nc	nc	nc
Purch R&D	nc	nc	nc	nc	nc	nc	nc	nc	nc	nc	nc	nc	nc	nc	nc
Int impair	nc	nc	nc	nc	nc	nc	nc	nc	nc	nc	nc	nc	nc	nc	nc
Debt ext	nc	nc	nc	nc	nc	nc	nc	nc	nc	nc	nc	nc	nc	nc	nc
Total	31	26	23	39	33	25	26	25	17	15	7	7	8	7	0

Panel C: Abbreviations Used for Line Items from the Tables in Accounting Trends and Techniques

Gains

Special Items
 Asset sale: Sale of assets
 Liab reversal: Liability accrual reduced
 Litigation: Litigation settlements
 Emp benefit: Employee benefit/pension related
 Debt ext: Debt extinguishment

Losses

Special Items
 Write-down: Write-down of assets
 Asset sale: Sale of assets
 Restructuring: Restructuring of operations
 Litigation: Litigation settlements
 Cleanup: Environmental cleanup
 PRD: Purchased R&D
 Intan impair: Impairment of intangibles
 Debt ext: Debt extinguishment

Not Tabulated:

Non-Special Items (per Compustat manual)

Interest

Equity in earnings of investees

Not Tabulated:

Non-Special Items (per Compustat manual)

Foreign currency transactions

Equity in losses of investees

(continued on next page)

Panel C: Abbreviations Used for Line Items from the Tables in *Accounting Trends and Techniques*

Dividends	Minority interests
Foreign currency transactions	Intangible asset amortization
Change in fair value of derivatives	Sale of receivables
Royalties	Royalties
Rentals	Merger costs: Merger costs
Insurance recoveries	Distributions of preferred securities of subsidiary trust
Public offering of subsidiary stock	Change in fair value of derivatives
Revenue of nonhomogeneous operations	Start-up costs
Sale of tax benefits	

"nc" Indicates this data were not compiled by *Accounting Trends and Techniques* and an asterisk (*) indicates that this item is less than 0.5 percent.

Panels A and B: Each cell reports the percent of firms out of the 600 surveyed by *Accounting Trends and Techniques* that reported the item in that row. Identification of Compustat special items is based on items specifically identified in the Compustat manual. However, additional items may be included if the firm identifies them as special items, so this list is a guide only, not determinative (see Standard & Poor's 2003).

V. ALTERNATIVE EXPLANATIONS FOR THE INCREASING FREQUENCY OF SPECIAL ITEMS

In general there are three potential reasons for the increase in special items over the past several decades. First, changes in economic activity could result in firms recording more special items than before. For example, it could be more common in the 2000s than in the 1960s for firms to experience impaired assets, sell an asset before it is used up, or lay off employees. Such changes in economic activity can lead to an increase in special items even when there has been no change in formal accounting standards. Second, new accounting standards could either require firms to record a special item that was previously ignored or provide more specific guidance that effectively encourages firms to record a special item more frequently than before. Below, we discuss specific accounting standards that relate to special items. Third, changes in accounting and disclosure practice can lead to firms recording more special items than previously even when economic events and formal accounting standards are not changing. For example, firms may begin to more actively write down assets that have been impaired rather than continuing to depreciate them according to the original depreciation schedule.¹² Such changes in accounting practice could occur if the financial reporting incentives for the firm or manager change. However, these changes are difficult to disentangle from changes in accounting practice that are precipitated by changes in economic activity or formal accounting standards.¹³

Despite a significant amount of research on special items, there is little evidence on which of these three potential reasons for the increase in special items is most important. Indeed, Frankel (2009, 239) states that “we do not understand why the use of special items has increased.” In this section, we provide evidence on these potential reasons for the increase in special items in recent decades. Because changes in accounting practice are difficult to separate from the other two reasons, we focus on economic activity and accounting standards as two alternative, non-exclusive explanations for the increase in special items and the resulting effect on the properties of accounting earnings.

DT also consider accounting standards and economic explanations for their findings by testing whether their main results obtain for both accrual and cash flow expense measures. They find that the changes in the properties of earnings are attributable primarily to accruals and conclude that “accounting-based factors are a substantial and perhaps even primary determinant of the documented temporal patterns of earnings characteristics, while real-economy changes play a secondary role” (DT 2008, 1453).¹⁴ Our finding that special items are responsible for the change in earnings properties is consistent with DT’s finding with respect to accruals because special items are almost exclusively accruals (Dechow and Ge 2006). However, it is not clear *ex ante* whether more frequent accruals for restructuring charges, impairments, and asset sales are driven predominantly by new accounting standards or by economic factors.

¹² Alternatively, firms may begin to more transparently disclose write-downs that they were already recording but not separately disclosing. We note, however, that more transparent disclosure of a constant frequency and magnitude of special items would not explain the changes from the early to the late period reported in Table 1.

¹³ Recent research by McVay (2006) and Fan et al. (2010) provides evidence that managers opportunistically manage recurring firm performance by including recurring charges as one-time special items. However, neither of these studies provides evidence on whether such classification shifting is either increasing over time or affects the frequency of special items. Moreover, such classification shifting, which results in special items that are more persistent, is likely to partially mitigate the disruptive effect of special items on the revenue-expense relation documented in Tables 2 and 3. Thus, while classification shifting is an important issue in accounting for special items, it is not likely to contribute to the role of special items in the revenue-expense relation.

¹⁴ DT also examine whether changes in industry composition in their sample drive their results as another potential economic explanation. We find that special items are increasing across virtually all industries (untabulated).

The Role of Economic Events

In this subsection we examine whether an increase in economic activity associated with special items has contributed to the increasing frequency of special items. To examine the role of changes in economic activity, we construct an index that includes specific economic events that cannot arise from bookkeeping practices alone, but are frequently associated with firms recording special items. We then examine whether this index is associated with the frequency of special items across firms and over time.

We begin by identifying five firm-year indicator variables that represent real economic events that are likely to be associated with the most common special items in Table 4. Our first measure (*Emp*) takes the value of 1 if the firm has experienced negative employee growth that year, and 0 otherwise. Negative employee growth is often the result of contracting business, and can be the direct result of restructuring layoffs that result in a special item. Our second and third measures are merger and acquisition activity (*M&A*) and discontinued operations (*Disc*). These variables reflect underlying growth or contraction through firm entry and exit in markets. Firms that enter or exit new markets or product lines are also likely to restructure operations or sell or write off assets, so we expect *M&A* and *Disc* to be related to the incidence of special items. The fourth and fifth measures indicate if the firm has an operating loss (*Loss*) or if the firm has negative sales growth (*Grow*). Firms that experience an operating loss or declining sales often write off assets or restructure their businesses to become more competitive, so these variables are also expected to be associated with the probability of recording a special item.

To provide evidence that these economic event variables are linked to the incidence of special items, we combine them into a single measure and examine the relationship of this measure to the incidence of special items. We give each observation a point if the observation has a 1 for each indicator variable, and then sum the points across the five economic event variables for each firm-year observation and label that sum *E-Score*. Thus, firm-year observations with an *E-Score* of 0 (5) have a 0 (1) for each indicator variable, indicating a low (high) level of economic activity associated with special items.

To successfully capture economic events associated with the increasing incidence of special items, this index should increase over time and should also be associated with the cross-sectional probability of recording a special item. To examine whether the index is increasing over time, we calculate the average *E-Score* each year during the study period and then average the annual averages for four sub-periods, 1967–1974, 1975–1984, 1985–1996, and 1997–2005. The results are reported in the first row of Table 5, Panel A. The average *E-Score* increases steadily from 0.73 for the first sub-period to 1.15 for the last sub-period, an increase of nearly 60 percent. The correlation between average *E-Score* and a time trend is 0.55 ($p < 0.01$).¹⁵

To provide further evidence that *E-Score* is increasing over time, we divide the sample each year into three groups on the basis of *E-Score*. The “high” group consists of observations with *E-Score* equal to 3, 4, or 5, the “middle” group consists of observations with *E-Score* equal to 1 or 2, and the “low” group consists of observations with *E-Score* equal to 0. The next two rows in Panel A report that the middle (high) *E-Score* group has 50 (4) percent of the observations in the first sub-period, rising to 58 (11) percent of the observations in the last sub-period, confirming the increase in *E-Score* over time.

The results for each year are presented graphically in Figure 2, along with fitted trend lines, which indicate that the incidence of large special items for the high *E-Score* group is much higher than for the low group. Further, the incidence for the high group is growing at more than twice the rate for the low group. The trend line slope for the high group is 0.017 (i.e., an average increase

¹⁵ We also observe statistically significant (at least $p < 0.10$) increases over time for all five components of *E-Score*.

TABLE 5
***E-Score* Classifications and Special Items**

Panel A: *E-Score* over Time

	Mean 67–74	Mean 75–84	Mean 85–96	Mean 97–05	Time Trend	p-value
Level of <i>E-Score</i>	0.73	0.90	1.03	1.15	0.55	<0.01
Incidence of Middle <i>E-Score</i> Observations	0.50	0.52	0.54	0.58	0.49	<0.01
Incidence of High <i>E-Score</i> Observations	0.04	0.07	0.10	0.11	0.54	<0.01

Panel B: Cross-Sectional Relation between *E-Score* and Special Item Incidence

	Low <i>E-Score</i>	Middle <i>E-Score</i>	High <i>E-Score</i>	Diff
1967–1974	0.7	1.9	5.1	4.4
1975–1984	5.0	10.8	20.1	15.1
1985–1996	13.4	26.6	44.7	31.2
1997–2005	22.6	38.0	55.7	33.2
Diff	21.9	36.1	50.7	

We rank each firm-year observation on the incidence of five economic event variables from 1967 to 2005: negative employee growth, M&A, discontinued operations, negative revenue growth, operating losses. We give a firm-year observation one point for each indicator variable coded as 1. We then sum the points across the five variables for each firm-year observation, and label that sum *E-Score*. Thus, across observations, the *E-Score* can conceivably vary from 0 points (indicating the presence of no economic events) to 5 points (indicating the presence of all economic events). We then divide the sample each year into three groups: 0 points for the “Low” group (roughly bottom 25 percent of sample), 1 to 2 points for the “Middle” Group, and 3–5 points for the “High” group (roughly top 10 percent of sample).

In Panel A, we plot mean values of *E-Score* over four sub-periods between 1967–2005, along with a correlation with this time-series and a trend variable.

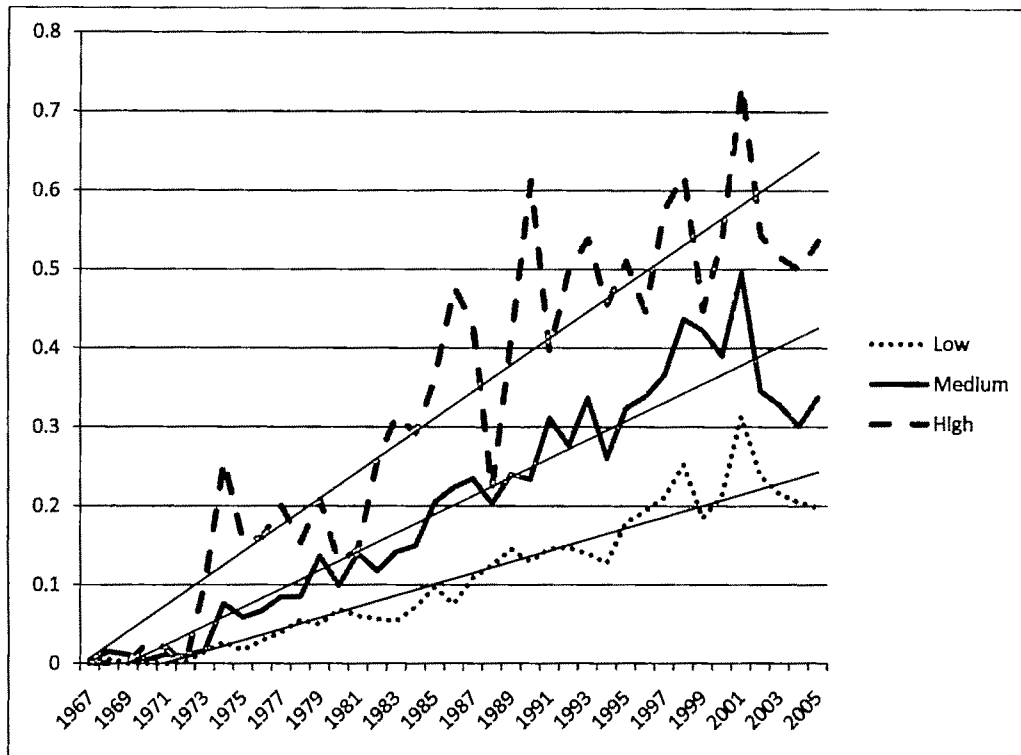
In Panel B, we compute the incidence of large special items for each of the three groups each year and then average over the four sub-periods. Bolded differences are significant at the 1 percent level (two-sided).

in the incidence of special items of 1.7 percent per year), compared with a slope of 0.007 for the low group, and the difference is significant at $p < 0.01$ (untabulated). This simple analysis provides strong evidence that our index of economic events is increasing over time and that it captures cross-sectional variation in the incidence of special items.

To provide further evidence that *E-Score* is increasing over time and is associated with the cross-sectional probability of recording a special item, we compute the average incidence of special items for the three *E-Score* groups for each of four sub-periods. The results, reported in Table 5, Panel B, reveal two clear patterns. First, the rows indicate that the average incidence of large special items for the high *E-Score* firms is more than double the incidence for the low *E-Score* firms in all four sub-periods. For example, in the most recent period, from 1997 to 2005, the incidence of large special items for firms in the high *E-Score* group is 55.7 percent, as compared to only 22.6 percent for the firms in the low *E-Score* group. These patterns indicate that *E-Score* is positively associated with the incidence of special items.¹⁶

¹⁶ To document that variation in *E-Score* is directly related to deterioration in the revenue-expense relation, we also re-estimate Equation (1) each year separately for each of the three *E-Score* groups. In untabulated results we find that the coefficient estimate on contemporaneous expense declines overall as *E-Score* increases: 0.96, 0.98, and 0.84 for the low, middle and high *E-Score* groups, respectively.

FIGURE 2
Firm-Specific Economic Variables and Special Items



This figure graphs the incidence of special items for firms based on *E-Score* by year. *E-Score* is defined in Table 5. Firms in the “High” group have a score of 3 or higher, firms in the “Medium” group have a score of 1 or 2, while firms in the “Low” group have a score of 0. Scores are calculated by ranking each firm-year observation on each of the five economic variables in Table 5 from 1967 to 2005, and giving the observation one point for each indicator variable. We then sum the points across the five variables for each firm-year observation.

Second, the columns indicate that the incidence of special items increases through time for all three groups, and the differences between the incidence for the early sub-period and the late sub-period are significant at the 0.01 level for all three groups (last row). Thus, special items become more common even for relatively fixed levels of our economic activity index, indicating that *E-Score* does not capture all determinants of special items. This is reinforced by the variation around the fitted lines in Figure 2. One explanation for this finding is that changes in economic activity not captured by the five simple variables comprising *E-Score* are responsible for increases in special items. An alternative explanation is that, holding economic events related to special items constant, other factors, such as new accounting standards, led firms to record more special items. We examine this explanation in the next subsection.

The Role of Accounting Standards

To examine the role of accounting standards, we identify the specific standards that were adopted during our study period that are most likely to have had an effect on the frequency of

special items. We then examine the frequency of special items both in the years (1) prior to implementation of these standards (when they had no effect,) and (2) after implementation of these standards (when the effect, if any, should take place).

The general accounting standard that applied to special items for most of the study period is APB Opinion No. 30 (AICPA 1973) on reporting the results of operations, which became effective in 1973. This standard does not use the term “special items” but rather paragraph 26 states that “A material event or transaction that is unusual in nature or occurs infrequently but not both ... should be reported as a separate component of income from continuing operations. The nature and financial effects of each event or transaction should be disclosed on the face of the income statement or, alternatively, in notes to the financial statements.” The results in Table 4 suggest that five items account for the majority of both the increase and the level of these separately disclosed income components: asset write-downs, sales of assets, restructurings, and, toward the end of the period, intangibles impairment and debt extinguishment. We examine specific accounting standards that were adopted during the study period that apply to each of these items.

Our search reveals that the earliest accounting standard related to special items adopted during our study period is Staff Accounting Bulletin 67 on restructurings, which was issued by the SEC in 1986 and is discussed in more detail below. However, Figure 1 suggests that the frequency of special items was increasing prior to 1986. To evaluate whether this is the case, we estimate the following time-series regression, with one observation per year from 1975 to 1985:

$$\%SI_t = \gamma_0 + \gamma_1 Trend_t + \gamma_2 E-ScoreAvg_t + \varepsilon_t, \quad (4)$$

where $\%SI_t$ is the percent of firms reporting a special item in year t , and $Trend$ is a variable that goes from 1 (for 1975) to 11 (for 1985). $E-ScoreAvg_t$ is the average for $E-Score$ (defined earlier) across firms for year t . We estimate Equation (4) using *Accounting Trends and Techniques* data.¹⁷ We compute $\%SI$ as the sum of the percent of gain special items and loss special items reported in Panels A and B of Table 4 (e.g., for 1983 we use $27 + 26 = 53$ percent). As the first column of Table 6 demonstrates, the estimated coefficient for $Trend$ is 0.012 ($t = 7.67$; $p < 0.01$) and the estimated coefficient for $E-ScoreAvg$ is 0.032 ($t = 2.08$; $p = 0.071$). These results indicate a pronounced increase in total special items during the 11-year period prior to the first accounting standard related to special items, as well as some evidence of a role for our economic event proxy, $E-Score$.

Asset Write-Downs

In 1984, the EITF acknowledged that the frequency of asset write-downs had been increasing over time, despite no formal guidance in GAAP for their recognition (see EITF 84-28, in which no consensus on a standard was reached). The first standard to explicitly address asset write-downs was SFAS 121, effective in 1996, with early adoption in 1995 “encouraged.” This standard was superseded by SFAS 144 in 2002, but virtually all of the asset impairment rules were left intact. In Table 4, Panel A, asset write-downs increase from 11 percent in 1994 to 19 percent in 1996, but we observe an upward trend both before and after that interval, and larger single year increases in 1998 and 2001. Further, we note that SFAS 121 explicitly cites the widespread reporting of asset impairments without specific authoritative guidance as motivation for the standard (paragraph 5), suggesting this standard was at least partly reactive in nature (see Francis et al. 1996, 117–118).

To evaluate whether SFAS 121 increased asset write-downs, we use data from *Accounting*

¹⁷ Results are very similar and inferences are unchanged using Compustat data for this test. We report results utilizing the *Accounting Trends and Techniques* data for consistency throughout the tests reported in Table 6.

TABLE 6
Regression Analysis
Special Item Incidence and Accounting Standards

	Model 4	Model 5	Model 6	Model 7
Intercept	-0.090*** (-4.69)	-0.110*** (-3.44)	0.003 (0.08)	-0.352*** (-6.05)
Trend	0.012*** (7.67)	0.006*** (3.98)	0.007*** (5.92)	0.015*** (5.66)
E-ScoreAvg	0.032* (2.08)	0.069** (2.29)	0.101** (2.29)	0.201*** (5.19)
D86				-0.019 (-0.68)
D95		0.118*** (4.78)		-0.037 (-1.19)
Adj. R ²	0.914	0.910	0.688	0.887
n	11	30	30	26

*, **, *** Represent significance at the 10 percent, 5 percent, and 1 percent levels, respectively, in one-tailed tests (positive coefficients expected), except the intercept.

t-statistics are in parentheses below the coefficient estimates.

This table presents results from estimations of the following models:

$$\text{Model 4: } \%SI_t = \gamma_0 + \gamma_1 \text{Trend}_t + \gamma_2 \text{E-ScoreAvg}_t + \varepsilon_t.$$

$$\text{Model 5: } \%WD_t = \gamma_0 + \gamma_1 \text{Trend}_t + \gamma_2 D95_t + \gamma_3 \text{E-ScoreAvg}_t + \varepsilon_t.$$

$$\text{Model 6: } \%SA_t = \gamma_0 + \gamma_1 \text{Trend}_t + \gamma_2 \text{E-ScoreAvg}_t + \varepsilon_t.$$

$$\text{Model 7: } \%Rest_t = \gamma_0 + \gamma_1 \text{Trend}_t + \gamma_2 D86_t + \gamma_3 D95_t + \gamma_4 \text{E-ScoreAvg}_t + \varepsilon_t.$$

The models are OLS time-series regressions with 11–30 annual observations from 1975–2005, as noted under the model number, depending on data availability from *Accounting Trends and Techniques*, summarized in Table 4.

E-ScoreAvg is annual average of the E-Score variable defined in Table 5. All main variables are defined in the Appendix.

Trends and Techniques in Table 4 to estimate the following time-series regression, with one observation each year from 1976 to 2005:

$$\%WD_t = \gamma_0 + \gamma_1 \text{Trend}_t + \gamma_2 D95_t + \gamma_3 \text{E-ScoreAvg}_t + \varepsilon_t, \quad (5)$$

where $\%WD_t$ is the percent of firms surveyed by *Accounting Trends and Techniques* that reported an asset write-down in year t , *Trend* takes values from 1 (for 1976) to 30 (for 2005), and *D95* is an indicator variable that takes the value of 1 if the year is 1995 or later, and 0 otherwise. *E-ScoreAvg* _{t} is as defined earlier.

The second column of Table 6 reports estimates from Equation (5). Consistent with the anecdotal observations in EITF 84-28 and SFAS 121, there is evidence of a strong independent time trend in the incidence of asset write-downs ($t = 3.98$; $p < 0.01$). There is also evidence that our index of economic events helps explain the incidence of write-down activity over time ($t = 2.29$; $p = 0.03$). Finally, the coefficient for *D95* is 0.118 ($t = 4.78$; $p < 0.01$), indicating that the incidence of asset write-downs increased about 12 percentage points in the years following SFAS

121. Thus, it appears that SFAS 121 likely encouraged or legitimized asset write-down activity.

Sale of Assets

Table 4 reports a slow increase over time for this item for both gains and losses. However, because no specific standard applies to asset sales, this contribution does not appear to be due to changes in accounting standards. To evaluate whether the frequency of asset sales is increasing even though it is not affected by a new accounting standard, we estimate the following time-series regression, with one observation per year from 1975 to 2005:

$$\%SA_t = \gamma_0 + \gamma_1 Trend_t + \gamma_2 E-ScoreAvg_t + \varepsilon_t, \quad (6)$$

where $\%SA_t$ is the percent of firms surveyed by *Accounting Trends and Techniques* that reported a gain or loss from asset sales in year t , and $Trend$ is a variable that goes from 1 (for 1975) to 31 (for 2005). $E-ScoreAvg_t$ is as defined earlier. The third column in Table 6 indicates that the estimated coefficient for $Trend$ is 0.007 ($t = 5.92$; $p < 0.01$), consistent with the frequency of gains and losses from asset sales growing significantly during this period when no accounting standards related to asset sales were adopted. In addition, the estimated coefficient for $E-ScoreAvg$ is 0.101 ($t = 2.29$; $p = 0.03$), consistent with economic events playing a role in the increase in asset sales.

Restructuring of Operations

The first standard to address this issue is SEC Staff Accounting Bulletin 67, issued in 1986, which states that restructurings should be considered part of operations and disclosed separately if material. Later, EITF 94-3 (adopted in 1995) provides guidance on when a restructuring liability can be recorded. This guidance was revised by SFAS 146 in 2002 (effective in 2003). Restructurings were first reported in *Accounting Trends and Techniques* in 1980 when the frequency was 8 percent. Six years later, the SEC reacted to this new disclosure by issuing SAB 67. Overall, there is a steady increase in the frequency of restructurings throughout the period with spikes in 1986, 1993, 2001, and 2002. One of these is coincident with SAB 67, which may have had the effect of legitimizing this type of disclosure. The spike in 1993 could have resulted from the economic downturn at that time, and may have precipitated the Emerging Issues Task Force to consider the issue in 1994. The additional spikes in 2001 and 2002 are likely the result of 9/11 and the bursting of the technology stock bubble.

To evaluate whether the frequency of special items was affected by these changes in accounting standards, we estimate the following time-series regression, with one observation per year from 1980 to 2005:

$$\%Rest_t = \gamma_0 + \gamma_1 Trend_t + \gamma_2 D86_t + \gamma_3 D95_t + \gamma_4 E-ScoreAvg_t + \varepsilon_t, \quad (7)$$

where $\%Rest_t$ is the percent of firms surveyed by *Accounting Trends and Techniques* that reported a restructuring loss in year t , $Trend$ is a variable that goes from 1 (for 1980) to 26 (for 2005), and $D86$ ($D95$) is an indicator variable that takes the value of 1 for all years from 1986 (1995) and later, and 0 (0) otherwise. Thus, $D86$ captures the incremental increase in the incidence of restructuring activity from 1986 to 1994 (after SAB 67), and $D95$ captures the incremental shift in 1995 and thereafter (after EITF 94-3). $E-ScoreAvg_t$ is as defined earlier.

In the fourth column of Table 6, the estimated coefficient for $Trend$ is 0.015 ($t = 5.66$; $p < 0.01$), indicating a significant increase in the frequency of restructurings. The estimated coefficients on $D86$ and $D95$ are insignificant, indicating no increase in restructurings after these accounting standards. Finally, the estimated coefficient of 0.201 for $E-ScoreAvg$ is significant ($t = 5.19$; $p < 0.01$), consistent with economic activity explaining some of the increasing incidence of restructurings.

Impairment of Intangibles and Debt Extinguishment

Intangible assets were subject to SFAS 121, but *Accounting Trends and Techniques* did not begin reporting this item until 2003, after the adoption of SFAS 142. This standard applies specifically to intangible assets and took effect in 2002. SFAS 145, which took effect in 2003 for calendar-year firms, applies to debt extinguishment transactions. This item first appeared in *Accounting Trends and Techniques* in 2003 at the 10 percent frequency level, and has maintained about that level since. We do not conduct tests for these items because SFAS 142 and SFAS 145 were adopted too recently to explain the nearly 30 years of increasing frequency of special items prior to 2002.

The results summarized above indicate that the incidence of total special items was increasing for some years prior to the adoption of a specific accounting standard addressing special items. We also find that gains and losses from asset sales were increasing throughout the study period even though no new accounting standard affected these transactions. However, there is evidence that a new accounting standard for asset write-downs (SFAS 121) could have contributed to the increasing frequency of these special items. Importantly, our proxy for economic events (*E-ScoreAvg*) is significant in each test.¹⁸ Thus, in general, we observe a stronger and more sustained role for changes in economic activity as an explanation for the increasing frequency of special items and a narrower role for changes in specific accounting standards.

To combine the above analyses into one test, we use Compustat data to compare the relative contributions of specific accounting standards and the economic activity represented by *E-Score* to the increase in the incidence of *total* special items over time. We estimate the following pooled time-series and cross-sectional regression at the firm-year level for 1967–2005:

$$SI_{it} = \gamma_0 + \gamma_1 D86_{it} + \gamma_2 D95_{it} + \gamma_3 D96_{it} + \gamma_4 D02_{it} + \gamma_5 D03_{it} + \gamma_6 E-Score_{it} + \gamma_7 Trend_{it} + \varepsilon_{it} \quad (8)$$

In Equation (8), SI_{it} is an indicator variable equal to 1 if firm i in year t reported a large special item, and 0 otherwise. $D86$, $D95$, and $D96$ are defined as in Equations (4)–(7) and $D02$ and $D03$ are defined analogously (2002 for SFAS 142 and 2003 for SFAS 145 and SFAS 146). Thus, these indicator variables are defined such that successive variables indicate the incremental frequency of special items until the next indicator variable is turned on. For example, the coefficient estimate on $D96$ represents the incremental frequency for the period 1996 to 2001 (prior to $D02$), relative to the period prior to 1996. *E-Score* and *Trend* are as defined above.¹⁹

The results for our estimation of Equation (8) are presented in the first column of Table 7. None of the coefficient estimates for the accounting standard indicator variables are positive. In fact, several are significantly negative.²⁰ Thus, once the time trend and *E-Score* are controlled, these results provide no evidence that the increase in total special items is due to the implementation of these accounting standards. In contrast, the coefficient estimate on *E-Score* is positive (0.512) and highly significant ($t = 24.87$; $p < 0.01$). The coefficient estimate on *Trend* is also positive (0.127) and highly significant ($t = 7.98$; $p < 0.01$), indicating that, even after controlling for *E-Score*, the incidence of special items is increasing substantially during the study period.

As an alternative analysis, we also estimate the following time-series regression with one observation for each year from 1967–2005:

¹⁸ Inferences with respect to accounting standards and the time trend are unchanged if *E-ScoreAvg* is omitted.

¹⁹ We are grateful to an anonymous reviewer for suggesting the analysis in models 8 and 9.

²⁰ The coefficient estimates for $D02$ and $D03$ are significantly *negative*, indicating that special items fell significantly from the 2001 level (likely due to 9/11). The coefficient estimate on $D95$ is also significantly negative.

TABLE 7
Regression Analysis
Accounting Standards and Economic Events

	Model 8	Model 9
Intercept	-4.583*** (-20.12)	-0.073*** (-4.06)
D86	-0.173 (-1.03)	0.029 (1.51)
D95	-0.249** (-2.25)	0.005 (0.17)
D96	-0.129 (-1.64)	0.046 (1.51)
D02	-0.669*** (-8.25)	-0.062* (-2.03)
D03	-0.321*** (-4.93)	-0.032 (-0.97)
E-Score	0.512*** (24.87)	
E-ScoreAvg		0.076*** (3.53)
Trend	0.127*** (7.98)	0.008*** (7.03)
Pseudo/Adj. R ²	0.147	0.954

*, **, *** Represent significance at the 10 percent, 5 percent, and 1 percent levels, respectively, in one-tailed tests (positive coefficients expected), except the intercept.

t-statistics are in parentheses below the coefficient estimates and are based on standard errors clustered by firm and year.

This table presents results from estimations of the following models:

$$\text{Model 8: } SI_{i,t} = \gamma_0 + \gamma_1 D86_{i,t} + \gamma_2 D95_{i,t} + \gamma_3 D96_{i,t} + \gamma_4 D02_{i,t} + \gamma_5 D03_{i,t} + \gamma_6 E\text{-Score}_{i,t} + \gamma_7 Trend_{i,t} + \varepsilon_{i,t}.$$

$$\text{Model 9: } \%SI_t = \gamma_0 + \gamma_1 D86_t + \gamma_2 D95_t + \gamma_3 D96_t + \gamma_4 D02_t + \gamma_5 D03_t + \gamma_6 E\text{-Score}_t + \gamma_7 Trend_t + \varepsilon_t.$$

Model 8 is a pooled logistic regression at the firm-year level where *SI* is an indicator variable taking the value of 1 if a large special item is reported, 0 otherwise. *E-Score* is the firm-year variable defined in Table 5. The time period is 1967–2005.

Model 9 is OLS time-series regression with 39 annual observations from 1967–2005. In model 9, *E-ScoreAvg* is the annual average of *E-Score*.

Other variables are defined in the Appendix.

$$\%SI_t = \gamma_0 + \gamma_1 D86_t + \gamma_2 D95_t + \gamma_3 D96_t + \gamma_4 D02_t + \gamma_5 D03_t + \gamma_6 E\text{-ScoreAvg}_t + \gamma_7 Trend_t + \varepsilon_t \quad (9)$$

In Equation (9), $\%SI_t$ is the percent of observations with a special item greater than 1 percent of total assets for year *t*. The year indicator variables in Equation (9) are identical to those in Equation (8), while *E-ScoreAvg* and *Trend* are as defined earlier in Equations (4)–(7).

The results from estimating Equation (9) are presented in the second column of Table 7. Consistent with the results for Equation (8), none of the coefficient estimates on the accounting standard indicator variables are significantly positive. This suggests that these accounting standards did not have a significant effect on the incidence of total special items once the time trend

and *E-ScoreAvg* are controlled. The coefficient estimate on *E-ScoreAvg* is positive (0.076) and significant ($t = 3.53$; $p < 0.01$). This indicates that, after controlling for the time trend, the incidence of special items is 7.6 percent higher for every one-point increase in average *E-Score*.²¹

The results for Equations (8) and (9) provide no support that adoption of the specific accounting standards we investigate contributed significantly to the increase in total special items. In contrast, our composite measure of economic events is significantly related to the incidence of special items over the study period.²²

VI. COMPETITIVE PRESSURE AND THE CHANGING ECONOMIC ENVIRONMENT

The evidence presented in the previous sections suggests that the increasing incidence of special items in recent decades is attributable, at least in part, to changes in the frequency of economic events associated with recording special items. In this section, we examine whether the increase in this type of economic activity is related to increases in the level of competition in the U.S. (and global) economy over the last 40 years that are documented in previous research.

Prior Research on Increases in Competition

A growing literature provides evidence that the level of competition in the U.S. economy has increased over the last several decades due to changes in trade, technology, and regulation, among other factors. For example, Shepherd (1982) documents a large increase in the level of competition throughout the U.S. economy between 1958 and 1980, which he attributes primarily to antitrust policies, with additional contributions from import competition and deregulation. Pryor (1994) reports that U.S. concentration ratios, a common inverse indicator of competition used in the industrial organization literature, declined between 1958 and 1982 when adjusted for imports, indicating that imports have contributed to increasing competition in U.S. industries.²³

More recent studies by Comin and Philippon (2005), Gaspar and Massa (2006), and Irvine and Pontiff (2009) document increases in competition later in our study period that they link to increases in firm volatility or performance instability. In addition, Klein and Marquardt (2006) show that the well-documented increasing incidence of accounting losses is attributable mostly to nonaccounting factors such as real firm performance, consistent with an increase in competition over time. In the executive labor market, Huson et al. (2001) provide evidence that forced CEO turnover has increased in recent years, and DeFond and Park (1999) report that the frequency of CEO turnover is positively related to the extent of competition in the industry.

Additional studies provide evidence that firms take actions commonly associated with recording special items in response to competitive pressure. For example, Bernard et al. (2006) report that import-related competition is negatively associated with plant survival and growth, suggesting an association between import competition and asset write-offs. Holmstrom and Kaplan (2001, 121) argue that corporations now regularly transform themselves, through “large waves of merger, takeover and restructuring” that they attribute to “increased dominance of capital markets.” Harford (2005) provides evidence that merger waves between 1981 and 2000 were related to deregu-

²¹ We also estimate a “changes” version of Equation (9), where the indicator variables for accounting standards include both the implementation year and the following year to capture incremental changes in incidence following the passage of various accounting standards. The change in the percentage of large special items is strongly related to the change in average *E-Score* ($\Delta E\text{-ScoreAvg}$). The coefficient estimate for $\Delta E\text{-ScoreAvg}$ is 0.072 ($t = 2.86$), the coefficient for time trend is insignificant, and no accounting standard variable is positive and significant.

²² We view Equations (8) and (9) as complementary. Equation (8) allows *E-Score* to explain cross-sectional variation in the incidence of special items. This may control for “noise” unrelated to new accounting standards, resulting in more powerful tests for the effects of new accounting standards. Alternatively, this may control some of the variation due to new accounting standards, resulting in weaker tests for the effects of accounting standards.

²³ For additional discussion of increasing competition and its possible drivers, see Blinder (2000), Whitman (1999), Reich (2007), and London (2004).

lation and competition among industry peers (see Table 2). Taken together, this research provides direct evidence that competition has increased in the U.S. economy and indirect evidence that this increase in competition is related to company actions that are associated with recognizing special items.

Empirical Analysis

In this subsection, we examine whether increases in competitive pressure can help explain increases in our index of economic events related to special item recognition. Our first measure of competition is the Herfindahl index, a standard measure in industrial organization (e.g., Gaspar and Massa 2006). To compute the Herfindahl Index (*Herf*) we use the Compustat universe to sum across each industry and year each firm's squared market share, where market share is firm revenue divided by industry revenue. We supplement this standard measure of competition with Altman's (1968) z-score (*Bank*), a common measure of bankruptcy risk or financial distress used in accounting and finance research. We view financial distress as a logical byproduct of increased competition where the economic environment is less forgiving of operational errors. We measure z-score at the industry level in the prior fiscal year to better reflect the competitive environment the firm faces rather than the firm's individual financial distress. Measuring z-score at the industry level on a lagged basis also avoids a mechanical relation between components of *E-Score* and the accounting inputs used to calculate z-score.²⁴

We examine these two measures individually and in combination. To combine them, we compute a *P-Score* analogous to our *E-Score* by giving a point for each of the two competitive pressure variables above the median value for that variable based on the sample distribution from 1967–1974, and summing the points for each observation. This sum varies from 0 points (below the median for both variables) to two points (above the median for both variables).

If the economic activity captured by *E-Score* is associated with increasing competitive pressure, we expect our measures of competitive pressure to increase over time and to be correlated with *E-Score* in the cross-section. To examine the behavior of our pressure variables over time, we compute the annual average values for *Herf*, *Bank*, and *P-Score* for each of the four sub-periods used previously, as well as the correlation with a time trend. The results are reported in Table 8, Panel A. All three measures are increasing over time (becoming less negative for *Herf* and *Bank*) and all correlations with the time trend are positive and significant ($p < 0.01$).

To provide further evidence on this increasing trend, we divide the sample into three groups on the basis of *P-Score*. The “high” group consists of observations with *P-Score* equal to 2, the “middle” group consists of observations with *P-Score* equal to 1, and the “low” group consists of observations with *P-Score* equal to 0. The last row of Panel A reports that the average incidence of observations in the high *P-Score* group increases from 19 percent in the first sub-period to 45 percent in the last sub-period. Taken together, these results indicate that the annual averages for these pressure variables are increasing during our study period.

To examine whether our measures of competitive pressure are associated with *E-Score* in the cross-section, we compute the average *E-Score* for the three *P-Score* groups for each of the four sub-periods. The results are reported in Table 8, Panel B. The clearest pattern is looking across the columns for each row, where we observe that the average *E-Score* is increasing across *P-Score* groups (i.e., increasing in competitive pressure) in all four sub-periods. These increases are statistically significant for the last three sub-periods ($p < 0.01$). For example, in the period 1997 to 2005 the average *E-Score* is 0.80 for the low *P-Score* group, as compared to 1.20 for the high *P-Score* group.

²⁴ For both *Herf* and *Bank* we assign observations to industries based on SIC codes at the two-digit level.

TABLE 8
Levels of Competitive Pressure Variables over Time

Panel A: Competitive Pressure Variables by Time Period and Correlation with Time Trend

	Mean 67–74	Mean 75–84	Mean 85–96	Mean 97–05	Time Trend	p-value
Level of <i>Herf</i>	–0.07	–0.07	–0.06	–0.05	0.70	<0.01
Level of <i>Bank</i>	–3.97	–3.56	–3.41	–3.13	0.56	<0.01
Level of <i>P-Score</i>	0.95	1.15	1.17	1.40	0.65	<0.01
Incidence of High <i>P-Score</i> Observations	0.19	0.27	0.30	0.45	0.63	<0.01

Panel B: Average *E-Score* for *P-Score* Groups

	Low <i>P-Score</i>	Middle <i>P-Score</i>	High <i>P-Score</i>	Diff
1967–1974	0.69	0.73	0.74	0.04
1975–1984	0.78	0.89	0.96	0.18
1985–1996	0.86	1.06	1.03	0.17
1997–2005	0.80	1.14	1.20	0.40
Diff	0.11	0.41	0.46	

We use two variables to proxy for competitive pressure, the Herfindahl index (*Herf*) and Altman's z-score (*Bank*). See the Appendix for variable definitions. We give each observation in any period one point for each value of *Herf* or *Bank* below the median value for these variables in the period 1967–1974, and then sum points across observations to form an index called *P-Score*. This index can range from 0 to 2 points. We label observations with a value of 0 (roughly bottom 10 percent of the sample) as having a “Low” *P-Score*, observations with a value of 1 as having a “Middle” *P-Score*, and observations with 2 points as having a “High” *P-Score* (roughly top 25 percent of sample).

Panel A plots average levels of *P-Score* and its components over four sub-periods from 1967–2005, along with a correlation of these time-series with a trend variable. For *Herf* and *Bank*, we use cross-sectional medians each year due to the strong skewness of these continuous variables. We also tabulate the incidence of “High” *P-Score* observations over time. In Panel B, we calculate average values of *E-Score* across the three *P-Score* groups each sub-period. Bolded differences are significant at the 1 percent level (two-sided).

Looking down the columns in Panel B of Table 8, the results also reveal a statistically significant increase in average *E-Score* over time for observations only in the middle and high *P-Score* groups. This indicates that economic events associated with special items have not increased among firms facing low levels of competitive pressure, but have increased for firms facing medium or high levels of economic pressure.²⁵ In addition, we re-estimate model 8 after adding *P-Score*. In this regression, *P-Score* is positive and significant, suggesting that competitive pressure contributes to the incidence of special items (untabulated). Overall, this evidence indicates that increases in economic events linked to special items are related to increasing competitive pressure in the economy.

²⁵ In untabulated analysis, we also estimate annual regressions of *E-Score* first on the two economic pressure variables, and then separately on *P-Score*. The averages for the estimated annual slope coefficients are positive and highly significant ($p < 0.01$) for both the components of *P-Score* and for *P-Score* itself.

VII. CONCLUSION

This study examines the reasons for the substantial decline in the contemporaneous association between revenue and expense over time, as documented by Dichev and Tang (2008). To isolate the source of this decline, we first disaggregate total expense into six components. We find that this decline (and the associated increase in earnings volatility and decrease in earnings persistence) is attributable primarily to a steady increase in the frequency of large special items.

We examine two potential reasons for this increase in special items. The first is that changes in specific economic activity have led to more special items. The second is that changes in specific accounting standards require (or encourage) firms to report more special items. We find little evidence that the specific accounting standards we study are a primary reason for the increase in special items. In contrast, we find much stronger evidence that changes in economic activity are a primary reason for the increase in special items. In particular, we find that an index of five indicators of specific economic events associated with recording special items has increased steadily over time and that high values for the index have more than twice the incidence of special items as low index values.

Finally, we examine whether increasing competition in the U.S. economy is related to increases in our index of economic events. We identify two indicators of competitive pressure at the industry level, finding that, both individually and when combined, they are increasing over time and are correlated with our economic activity index in the cross-section.

Our findings demonstrate the importance of disaggregating special items, and are important for standard-setters, given the frequency of these items. This is especially true in light of the recent discussion paper by the FASB and the IASB, *Preliminary Views on Financial Statement Presentation* (FASB 2008), which directly addresses this issue. The FASB's preliminary view is that special items should be disclosed in a "memo" column of the proposed reconciliation between cash flows and comprehensive income (paras. 4.48 to 4.52), while the IASB does not support this additional disclosure (para. 4.53).

While the evidence presented in this study suggests that changes in economic activity are a primary reason for the increase in special items and specific accounting standards contribute little to the increase in special items, our results do not prove the null hypothesis that accounting standards play no role in the increase in special items. The inability of our economic variables to fully explain the time trend in special items leaves open the possibility that accounting standards could explain some of this unexplained variation. Thus, it is possible that future research will develop stronger tests that are able to detect a role for accounting standards.

We also leave two additional important issues for future research. The first is the possibility that changes in financial reporting practice that are not related to either new accounting standards or to changes in economic activity have contributed to the increase in special items over time. For example, perhaps managers facing increasing capital market pressure have changed their accounting practices and are now more aggressively recognizing nonrecurring gains and losses, encouraging investors to ignore these items for performance evaluation. This issue is beyond the scope of this study. Also beyond the scope of this study is whether performance measures that so frequently include special items are "good accounting." While our evidence indicates that the increase in special items over time is linked to increases in economic events that affect firm welfare, we leave a more complete evaluation of the appropriateness of this accounting to future research.

APPENDIX DETAILED VARIABLE DEFINITIONS

Earnings Variables and Components

- Revenue (*Rev*) = sales (data12)/average total assets (data6);
- Expense (*Exp*) = sales less income before extraordinary items (data18)/average total assets (data6);
- Cost of Goods Sold (*COGS*) = data 41/average total assets (data6);
- Selling, General, and Admin (*SGA*) = data189/average total assets (data6);
- Depreciation (*DEPR*) = data 14/average total assets (data6), equals 0 if missing;
- Tax Expense (*TAX*) = data16/average total assets (data6);
- Special Items (*SI*) = $(-1) * \text{data17}/\text{average total assets (data6)}$ for decomposition analysis, equals 0 if missing. An indicator variable equal to 1 if the absolute value of special items is greater than or equal to 1 percent of average total assets (data6), 0 otherwise. Cross-sectional average measures annual incidence;
- Operating Income after Depreciation = data178/average total assets (data6);
- Other Expenses (*OTH*) = total expense less *COGS*, *SGA*, *DEPR*, *TAX*, and *SI*; and
- $\%SI_t$ = percent of sample firm-year observations with a special item greater than 1 percent of beginning total assets that year.

Economic Event Variables

- Negative Employee Growth (*Emp*) = indicator variable that takes the value of 1 if the firm had negative employee (data29) growth for the year relative to the preceding year;
- Merger and Acquisition Activity (*MA*) = indicator variable that takes the value of 1 if the firm engaged in a merger or acquisition during the fiscal year (from AFTNT1);
- Discontinued Operations (*Disc*) = indicator variable that takes the value of 1 if the firm had discontinued operations for the year (data66);
- Operating Loss (*Loss*) = indicator variable that takes the value of 1 if the firm had negative operating income after depreciation (data178) for the year;
- Negative Revenue Growth (*Grow*) = indicator variable that takes the value of 1 if the firm had negative revenue (data12) growth rate, 0 otherwise;
- E-Score* = sum of the five indicator variables *Emp*, *MA*, *Disc*, *Loss*, and *Grow*; and
- E-ScoreAvg* = cross-sectional average of *E-Score* for a given year.

Accounting Standard Variables

$D_{xx} = 1$ in 19xx or 20xx and beyond, and 0 prior to 19xx or 20xx.

Competitive Pressure Variables

Herfindahl Index (*Herf*) = sum for each industry-year of member firm's squared market share. Market share is firm revenue divided by industry revenue using all available Compustat data;

Altman's Bankruptcy Prediction

(*Bank*) = $(-1) \times$ the industry median z-score from Altman (1968) from the prior year, so that *Bank* is increasing in competition, varies from the lowest probability of bankruptcy (the lowest scores) to the highest probability of bankruptcy (the highest scores); and

P-Score = sum of a point for each of *Herf* and *Bank* that are above the sample median from 1967–1974.

REFERENCES

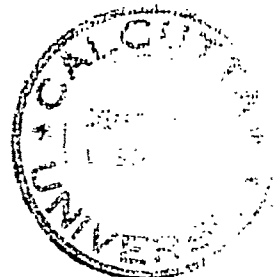
- Altman, E. I. 1968. Financial ratios, discriminant analysis, and the prediction of corporate bankruptcy. *The Journal of Finance* 23 (4): 589–609.
- American Institute of Certified Public Accountants. (AICPA). 1973. *Reporting the Results of Operations—Reporting the Effects of Disposal of a Segment of a Business, and Extraordinary, Unusual, and Infrequently Occurring Events and Transactions*. Accounting Principles Board Opinion No. 30. New York, NY: AICPA.
- Bernard, A., J. Jensen, and P. Schott. 2006. Survival of the best fit: Exposure to low-wage countries and the (uneven) growth of U.S. manufacturing plants. *Journal of International Economics* 68 (1): 219–237.
- Blinder, A. 2000. How the economy came to resemble the model. *Business Economics (Cleveland, Ohio)* (January): 16–25.
- Burgstahler, D., J. Jambalvo, and T. Shevlin. 2002. Do stock prices fully reflect the implications of special items for future earnings? *Journal of Accounting Research* 40 (3): 585–612.
- Collins, D., E. Maydew, and I. Weiss. 1997. Changes in the value-relevance of earnings and book values over the past forty years. *Journal of Accounting and Economics* 24 (1): 39–67.
- Comin, D., and T. Philippon. 2005. *The Rise in Firm-Level Volatility: Causes and Consequences*. Cambridge, MA: NBER Macroannuals.
- Cready, W., T. J. Lopez, and C. A. Sisneros. 2010. The persistence and market valuation of recurring nonrecurring items. *The Accounting Review* 85 (5): 1577–1615.
- Dechow, P. M. 1994. Accounting measures of earnings and cash flows as measures of firm performance: The role of accounting accruals. *Journal of Accounting and Economics* 18 (1): 3–42.
- , and W. Ge. 2006. The persistence of earnings and cash flows and the role of special items: Implications for the accrual anomaly. *Review of Accounting Studies* 11 (2–3): 253–296.
- , and C. M. Schrand. 2004. *Earnings Quality*. Charlottesville, VA: The Research Foundation of CFA Institute.
- DeFond, M., and C. Park. 1999. The effect of competition on CEO turnover. *Journal of Accounting and Economics* 27 (1): 35–56.
- Dichev, I., and V. Tang. 2008. Matching and the changing properties of accounting earnings over the last 40 years. *The Accounting Review* 83 (6): 1425–1460.
- Elliott, J., and D. Hanna. 1996. Repeated accounting write-offs and the information content of earnings. *Journal of Accounting Research* 34 (Supplement): 135–155.
- Fairfield, P., K. Kitching, and V. Tang. 2009. Are special items informative about future profit margins? *Review of Accounting Studies* 14 (2–3): 204–236.
- Fan, Y., A. Barua, W. Cready, and W. Thomas. 2010. Managing earnings using classification shifting: Evidence from quarterly special items. *The Accounting Review* 85 (4): 1303–1323.

- Financial Accounting Standards Board (FASB). 2008. *Preliminary Views on Financial Statement Presentation*. Norwalk, CT: FASB.
- Francis, J., D. Hanna, and L. Vincent. 1996. Causes and effects of discretionary asset write-offs. *Journal of Accounting Research* 34 (Supplement): 117–134.
- , R. LaFond, P. M. Olsson, and K. Schipper. 2004. Costs of equity and earnings attributes. *The Accounting Review* 79 (4): 967–1010.
- Frankel, R. 2009. Discussion of: Are special items informative about future profit margins? *Review of Accounting Studies* 14 (2–3): 237–245.
- Gaspar, J., and M. Massa. 2006. Idiosyncratic volatility and product market competition. *The Journal of Business* 79 (6): 3125–3152.
- Harford, J. 2005. What drives merger waves? *Review of Financial Studies* 77 (September): 529–560.
- Holmstrom, B., and S. N. Kaplan. 2001. Corporate governance activity in the United States: Making sense of the 1980s and 1990s. *The Journal of Economic Perspectives* 15 (2): 121–144.
- Huson, M., R. Parrino, and L. Starks. 2001. Internal monitoring mechanisms and CEO turnover: A long-term perspective. *The Journal of Finance* 56 (6): 2265–2297.
- Irvine, P., and J. Pontiff. 2009. Idiosyncratic return volatility, cash flows, and product market competition. *Review of Financial Studies* 22 (3): 1149–1177.
- Kee, H. L. 2009. The coefficient of the whole is an average of the coefficients of its parts: A special case of restricted least squares. *Economics Letters* 104 (3): 136–139.
- Klein, A., and C. Marquardt. 2006. Fundamentals of accounting losses. *The Accounting Review* 81 (1): 179–206.
- London, P. 2004. *Competition Solution: The Bipartisan Secret behind American Prosperity*. Washington, D.C.: AEI Press.
- McVay, S. 2006. Earnings management using classification shifting: An examination of core earnings and special items. *The Accounting Review* 81 (3): 501–532.
- Pryor, F. 1994. The evolution of competition in U.S. manufacturing. *Review of Industrial Organization* 9 (6): 695–715.
- Reich, R. 2007. *Supercapitalism: The Transformation of Business, Democracy, and Everyday Life*. New York, NY: Alfred A. Knopf.
- Shepherd, W. G. 1982. Causes of increased competition in the U.S. economy, 1939–80. *The Review of Economics and Statistics* 64 (4): 613–626.
- Standard & Poor's. 2003. *Compustat User's Guide*. Centennial, CO: The McGraw-Hill Companies, Inc.
- Whitman, M. N. 1999. *New World, New Rules: The Changing Role of the American Corporation*. Cambridge, MA: Harvard Business Press.

Internal Controls and Conditional Conservatism

Beng Wee Goh
Singapore Management University

Dan Li
Tsinghua University



ABSTRACT: This study examines the relation between internal controls and conditional conservatism ("conservatism"), also referred to as timely loss recognition. Using a sample of firms that disclose material weaknesses (MWs) in internal controls under the Sarbanes-Oxley Act (SOX), we find a positive relation between internal control quality and conservatism. Specifically, firms with MWs exhibit lower conservatism than firms without such weaknesses. Further, firms that disclose MWs and subsequently remediate these weaknesses exhibit greater conservatism than firms that continue to have MWs. Overall, these results are consistent with strong internal controls acting as a mechanism that facilitates conservatism. Our study contributes to the literature on the reporting effects of strong versus weak internal controls.

Keywords: *internal controls; conservatism; material weaknesses; Sarbanes-Oxley Act.*

Data Availability: *Data are available from sources identified in the text.*

I. INTRODUCTION

In 2002, the U.S. Congress passed the Sarbanes-Oxley Act (SOX) to improve the quality of financial reporting and to restore investor confidence in the reliability of financial statements. An important aspect of SOX is its internal control reporting requirements, which allow investors to be informed about the quality of a firm's internal controls. Specifically, Section 302 of the Act requires management to evaluate and report the effectiveness of disclosure and control procedures (SEC 2002), while Section 404 requires the auditor to form an opinion on the effective-

We thank Steve Kachelmeier (senior editor) and two anonymous referees for suggestions and guidance that have significantly improved this paper. We also thank Bryan Church, Arnold Schneider, Eugene Comiskey, Steve Matsunaga, Jayanthi Krishnan, Jeffrey Ng, Michael Willenborg, Patricia Dechow, Yuan Zhang, Zhaoyang Gu, Ole-Christian Hope, Stephen Penman, Chee Yeow Lim, Chih-Ying Chen, and participants at the 2008 AAA Annual Meeting for their comments. We also appreciate the financial support from Tsinghua University and funding through a research grant from the Office of Research, Singapore Management University.

Editor's note: Accepted by Steven Kachelmeier.

Submitted: January 2009
Accepted: October 2010
Published Online: April 2011

ness of internal controls (SEC 2003).¹ Using these internal control disclosures, studies have investigated whether the quality of internal controls affects accruals quality (Doyle et al. 2007a; Ashbaugh-Skaife et al. 2008), a firm's cost of capital (Ogneva et al. 2007; Beneish et al. 2008; Ashbaugh-Skaife et al. 2009), and the accuracy of management guidance (Feng et al. 2009). Given the importance of the internal control provisions as a means to improve the governance of firms, our study extends the literature on the reporting effects of strong versus weak internal controls by examining how the quality of internal controls is related to conservatism in financial reporting.

We focus on conservatism because it has been argued to provide several governance benefits, such as reducing agency conflicts and improving managerial investment decisions (Holthausen and Watts 2001; Watts 2003; Ball and Shivakumar 2005), enhancing the efficiency of debt contracts (Ahmed et al. 2002; Zhang 2008), and reducing litigation costs (Watts 2003). Following Basu (1997), we define conservatism as the higher degree of verification to recognize good news as gains than to recognize bad news as losses. That is, earnings reflect bad news more quickly than good news leading to timelier loss recognition. The literature generally refers to this as conditional conservatism (Beaver and Ryan 2005; Ball and Shivakumar 2005), which we simply refer to as "conservatism" in this study.²

Within the agency framework of positive accounting theory that recognizes the contracting benefits of conservatism, it is plausible to expect that strong internal controls facilitate conservative financial reporting. First, firms with a strong internal control environment (e.g., strong tone-at-the-top and/or good internal control culture) are more likely to understand the role of conservatism in contracting and in reducing agency conflicts. Consequently, these firms are more likely to favor its implementation. Further, to the extent that firms are committed to produce conservative reports, strong internal controls can facilitate this process by providing reliable accounting information. Weak internal controls can cause estimation errors in accounting numbers (Doyle et al. 2007a; Ashbaugh-Skaife et al. 2008), making the contractible variables less reliable for contracting and monitoring purposes. This can reduce the effectiveness of conservatism as a governance mechanism and, hence, the incentives to implement conservatism. Finally, weak internal controls could impede the timely recognition of losses, thereby leading to lower conservatism. We develop these arguments on the positive relation between internal control quality and conservatism more fully in Section II.

Moving away from the agency framework of positive accounting theory that argues the contracting benefits of conservatism, it becomes less clear why one would expect a positive relation between internal control quality and conservatism. First, it has been argued that conservatism can increase earnings management (Levitt 1998; Penman and Zhang 2002), which can exacerbate agency conflicts. Under this view, if strong internal controls serve as an effective monitoring mechanism that can mitigate agency problems (Jensen 1993), then we should instead expect strong internal controls to provide a disincentive for firms to produce conservative reports. Second, it is controversial whether conservatism is a desirable attribute of financial reporting. Barth (2008) argues that conservatism is not a qualitative characteristic under the conceptual framework of IASB, which specifies that accounting information should be unbiased. Conservatism, by promoting the deliberate understatement of book value and/or earnings, would render

¹ In Endnote 59 of the SEC's (2002) release on the *Final Rule: Certification of Disclosure in Companies' Quarterly and Annual Reports*, internal control is defined as "a process, effected by an entity's board of directors, management and other personnel, designed to provide reasonable assurance regarding the reliability of financial reporting."

² Another form of reporting conservatism, referred to as "unconditional conservatism," manifests itself in a systematic undervaluation of the firm's net assets. Ball and Shivakumar (2005) argue that unconditional conservatism cannot yield contracting efficiency because it can be observed and, hence, can be "undone" by users of the financial statements. Accordingly, we do not consider unconditional conservatism in our context.

financial statements not neutral. Under this investor value relevance framework that emphasizes that accounting numbers should be neutral inputs to investment decisions, the argument that strong internal controls would promote conservative accounting seems tenuous.

Given the above arguments and corresponding tension, we empirically test the relation between internal control quality and conservatism using firms' internal control disclosures under SOX. We obtain a sample of firms that disclosed at least one material weakness (MW) from January 2003 to November 2005 and deem these firms (hereafter referred to as "MW firms") to have low internal control quality due to the presence of MWs.³ We apply three measures of conservatism that are commonly used in the literature to capture the asymmetric timeliness in the recognition of economic losses: (1) the persistence of earnings changes measure in Basu (1997), (2) the accrual-based loss recognition measure in Ball and Shivakumar (2005), and (3) the timeliness of earnings to news measure in Basu (1997). To afford stronger inferences to be made about the effects of internal controls on conservatism, we further conduct intertemporal tests of changes in the status of internal controls. Specifically, we examine whether MW firms that subsequently remediate their weaknesses exhibit different levels of conservatism from MW firms that continue to have these weaknesses.

Results using all three measures of conservatism are consistent with a positive relation between internal control quality and conservatism. Specifically, we find that MW firms exhibit lower conservatism than control firms without such weaknesses. Further, we find that MW firms that subsequently remediate their weaknesses (i.e., show an improvement in internal control quality) exhibit greater conservatism than MW firms that continue to have these weaknesses. These results are robust to a battery of checks, which include controlling for the self-selection bias in the MW firms sample, controlling for the self-selection bias in the improvement of internal control quality, and using alternative conservatism measures suggested by Khan and Watts (2009) and Givoly and Hayn (2000). Taken together, our results are consistent with strong internal controls serving as a mechanism that facilitates conservatism.

Doyle et al. (2007a) and Ashbaugh-Skaife et al. (2008) find a positive relation between internal control quality and accruals quality, which is measured by the extent to which accruals are realized as cash flows or the size of abnormal accruals. Our study differs from their studies by examining conservatism, which is a form of accounting bias toward reporting low book values of stockholder equity, *conditional on firms experiencing contemporaneous economic losses* (Ball and Shivakumar 2005).⁴ In both Doyle et al. (2007a) and Ashbaugh-Skaife et al. (2008), the authors conjecture that their findings could be due to weak internal controls (1) causing unintentional (poor estimation ability) errors, or (2) failing to limit intentional income-increasing earnings management. However, Ashbaugh-Skaife et al. (2008) find that firms that report internal control weaknesses have significantly larger positive and larger negative abnormal accruals relative to

³ According to Auditing Standards No. 2 (PCAOB 2004), which was in effect during the time period of our study, an MW is "a significant deficiency, or combination of significant deficiencies, that results in more than a remote likelihood that a material misstatement of the financial statements will not be prevented or detected." A significant deficiency is "a control deficiency, or combination of control deficiencies, that adversely affects the company's ability to initiate, authorize, record, process, or report external financial data reliably in accordance with generally accepted accounting principles such that there is more than a remote likelihood that a misstatement of the company's annual or interim financial statements that is more than inconsequential will not be prevented or detected."

⁴ Ball and Shivakumar (2006) argue that conventional accruals models assume that the relation between accruals and cash flows are linear and do not take into account of the asymmetry in the gain and loss recognition role of accruals. Hence, these models misspecify the accounting accrual process and misestimate discretionary accruals, resulting in incorrect inferences about earnings quality. By examining the timely loss recognition property of earnings, our study takes into consideration the gain and loss recognition role of accruals and the nonlinearity implied by the asymmetry therein. This addresses the concerns in the use of the linear accrual models.

control firms, which suggests that the first factor is more plausible. Nonetheless, these two factors need not necessarily result in lower conservatism *per se*.

First, the unintentional estimation errors caused by weak internal controls should affect the recognition of bad news and good news in an unbiased and symmetrical manner. Accordingly, these estimation errors are not likely to cause conservatism *per se*, which results from the asymmetric response of earnings to bad versus good economic news. Second, even if weak internal controls fail to limit earnings management, the presence of income-increasing earnings management is not necessarily consistent with the presence of lower conservatism. Givoly et al. (2010, 221) argue that earnings management is situational or episodic, occurring when unmanaged earnings fail to meet an important reporting objective (e.g., meeting analysts' forecasts). Further, earnings management tends to be small-scale in nature due to the short supply of the needed positive accruals. Hence, if the company's accounting is generally conservative, earnings management will only temporarily interrupt this observed reporting pattern but will not obscure the presence of the more prevalent phenomenon of conservatism. The implication is that earnings management and conservatism can co-exist within a firm.⁵ Hence, our study complements Doyle et al. (2007a) and Ashbaugh-Skaife et al. (2008) by providing insights into how internal control quality is associated with the phenomenon of conservatism.

Overall, our study extends the literature on the reporting effects of strong versus weak internal controls (e.g., Doyle et al. 2007a; Ogneva et al. 2007; Ashbaugh-Skaife et al. 2008; Beneish et al. 2008). We show that strong internal controls can create an incentive for and facilitate conservative reporting. Our study also contributes to the literature that shows that conservatism is shaped, in part, by a firm's contracting and governance environment (e.g., Ball et al. 2000; Basu et al. 2001; Ball et al. 2003; Ball and Shivakumar 2005; Bushman and Piotroski 2006; Ahmed and Duellman 2007; LaFond and Roychowdhury 2008).

Section II develops the hypotheses, Section III explains the research design, Section IV presents the empirical results, Section V describes additional analyses, and Section VI concludes.

II. HYPOTHESIS DEVELOPMENT

In an attempt to restore investor confidence in firms' financial reporting, SOX requires firms to assess and disclose, and auditors to certify, the effectiveness of internal controls over financial reporting (SEC 2002, 2003). Presumably, regulators hope that these requirements can improve the quality of internal controls and enhance the reliability of financial reporting. Indeed, studies have shown that effective internal controls can enhance financial reporting quality, proxied by accruals quality and the size of abnormal accruals (Doyle et al. 2007a; Ashbaugh-Skaife et al. 2008). Further, Beneish et al. (2008) and Ashbaugh-Skaife et al. (2009) find that effective internal controls result in reduced information risk, which can lower a firm's cost of equity. However, Ogneva et al. (2007) fail to find a significant relation between internal control weaknesses and cost of equity after controlling for primitive firm characteristics and analyst forecast bias. Finally, Feng et al. (2009) find a positive relation between internal control quality and the accuracy of management guidance, consistent with ineffective internal controls causing errors in internal management reports. Our study extends this stream of research by examining how the quality of internal controls affects conservatism in financial reporting.

Conservatism has been argued to play an important governance role in mitigating agency conflicts and enhancing contracting efficiency within the firm (Holthausen and Watts 2001; Watts

⁵ Consistent with this line of reasoning, Givoly et al. (2010) find that public equity firms report more conservatively than their private equity counterparts, but at the same time, public equity firms have a greater propensity to manage income. The authors also conduct two tests to show that earnings management and conservative reporting may independently exist in their data.

2003; Ball and Shivakumar 2005). The reasoning is that conservatism reduces managers' ability and incentives to overstate earnings and net assets by requiring higher verification standards for gain recognition and reduces managers' ability to withhold information on expected losses. Thus, it prevents overcompensation of managers that is costly to recover *ex post* because of managers' limited liability and tenure.⁶ Similarly, in a debt-contracting setting, conservatism reduces managers' ability to loosen or avoid dividend restrictions and transfer wealth from bondholders to shareholders, thereby mitigating deadweight losses and increasing firm value (Ahmed et al. 2002; Zhang 2008). According to Ball and Shivakumar (2005), conservative accounting effectively limits the control rights of managers and transfers those rights back to the providers of finance earlier. Finally, Watts (2003) asserts that conservatism can also reduce a firm's expected litigation costs.

Existing studies suggest that the incentives for or against conservative reporting are shaped, in part, by a firm's contracting and governance environment (e.g., Ball et al. 2000; Basu et al. 2001; Ball 2001; Ball et al. 2003; Ball and Shivakumar 2005; Bushman and Piotroski 2006; Ahmed and Duellman 2007; LaFond and Roychowdhury 2008; LaFond and Watts 2008; Nikolaev 2010). For instance, Ahmed et al. (2002) argue that conservatism evolves as an efficient contracting mechanism to mitigate dividend policy conflicts between shareholders and bondholders. They find that the more severe such conflicts are (i.e., the higher the leverage), the more conservative the firm's accounting choices become. LaFond and Roychowdhury (2008) hypothesize and find that when managerial ownership is low, the severity of the agency problem increases, driving a demand for greater conservatism. Ball and Shivakumar (2005) find that private firms are less conservative than public firms and attribute this finding partly to the differences in governance structures and monitoring mechanisms between these two types of firms. Given that the internal control system plays an important governance role within the firm to monitor managers' behavior and to minimize agency costs (Fama and Jensen 1983; Jensen 1993), we conjecture a potential link between the quality of internal controls and conservative reporting.

Within the agency framework of positive accounting theory that recognizes the contracting benefits of conservatism, strong internal controls could therefore create an incentive for conservative accounting reports. According to the *Internal Control—Integrated Framework* (Committee of Sponsoring Organizations of the Treadway Commission 1992), the internal control environment is an important component of internal controls. We argue that firms with a strong tone-at-the-top (e.g., an effective board of directors that protects the interests of shareholders) and good internal control culture (e.g., a corporate culture that has a wealth-maximization orientation) are better able to understand the benefits of conservatism in reducing agency conflicts and deadweight loss, in contracting with shareholders and/or debtholders, and in reducing the litigation risks for directors, auditors, and managers. Consequently, these firms are more likely to favor the implementation of conservatism and promote an environment that emphasizes the adherence to conservative reporting.

To the extent that firms understand the benefits of conservatism and are committed to produce conservative reports, strong internal controls can facilitate the use of conservatism in the contracting process by providing reliable accounting information. This is because the financial reporting system, by emphasizing verifiable outcomes, supplies a rich set of variables that can be used for contracting purposes (Watts and Zimmerman 1986; Bushman and Smith 2001; Armstrong et al. 2010). For instance, financial covenants used in debt contracts to limit managerial discretion with

⁶ Conservative accounting reduces the tendency of managers with short-term horizons to invest in negative Net Present Value (NPV) projects, making managers aware that they will not be able to defer the recognition of losses to the future (Ball and Shivakumar 2005) and imposing greater costs to biasing financial reports upward (Guay and Verrecchia 2006).

respect to dividend distributions (Ahmed et al. 2002) are often defined in terms of accounting numbers. Similarly, managers' compensation can be tied to changes in book value (or earnings).

As mentioned earlier, weak internal controls can cause unintentional errors in accrual estimations to occur and impact the reported book value or earnings. For instance, the failure to have appropriate reconciliations and reviews in place allow for procedural errors in accrual estimations (Doyle et al. 2007a), and the lack of adequate policies, training, or diligence by company employees can affect the noise and magnitude of abnormal accruals (Ashbaugh-Skaife et al. 2008). Although these estimation errors cannot cause lower conservatism per se, they can make the contractible variables noisy and less reliable, providing a poor signal to the board of directors that conservatism has been used effectively in contracting and reducing the usefulness of conservatism in monitoring managers.⁷ In sum, by enhancing the ability to use accounting numbers effectively in contracts, strong internal controls can create an incentive for firms to use conservatism as a governance mechanism.

Finally, weak internal controls could also impede the timely recognition of losses, thereby leading to lower conservatism. For instance, the lack of proper policies and procedures to regularly value a firm's inventory or fixed assets prevents asset impairment from being discovered early and, hence, delays the recognition of losses. Unqualified accounting staff may lack the expertise to estimate the future cash flows of assets such as goodwill and fixed assets, and identify any reduction in future cash flows that would trigger early recognition of impairment losses. Weak internal controls in the monitoring of the firm's investment decisions can also provide the incentive for managers to avoid timely loss recognition in investment projects. One example is the lack of a proper process to provide the board of directors with timely information on the firm's investments for oversight and review, which could make it easier for managers to invest in or avoid divesting negative NPV projects. Further, control weaknesses over the computation of the present value of future profits of acquired businesses can result in potential loss-making projects not being discovered and terminated early. The Appendix provides some examples of MWs disclosed by firms that could potentially impede the timely recognition of losses.

The above arguments suggest a positive effect of conservative accounting reports in terms of contracting efficiency. However, moving away from the agency framework of conservatism as a beneficial governance mechanism, it becomes less clear why one would expect a positive link between internal controls and conservatism. First of all, it is questionable whether conservatism is beneficial to firms and their shareholders. For instance, Penman and Zhang (2002) contend that conservative accounting can create unrecorded reserves that provide managers with the flexibility to report more income in the future. The increased ability to manage earnings can adversely affect the predictive ability of current earnings for future earnings. In addition, three of the five financial reporting problems highlighted by former SEC Chair Arthur Levitt deal with the understatement of assets—big bath charges, creative acquisition accounting, and cookie jar reserves (Levitt 1998). If so, then conservative accounting reporters may actually have greater ability to smooth or manipulate earnings, which could exacerbate, rather than mitigate, agency problems. If strong internal controls serve as an effective monitoring mechanism to reduce agency conflicts (Jensen 1993),

⁷ One may question why a firm would not ensure that internal controls are effective or remediate internal control weaknesses promptly if the firm views strong internal controls as important in ensuring conservative reports. This could be due to several reasons. First, the firm may fail to constantly invest in internal controls to keep up with the growth of the firm. Second, managers may not be willing to invest in the time and resources to ensure effective internal controls because such actions divert attention away from the core business. Third, firms may face strained financial resources. Fourth, firms may have complex operations, such as multiple operating segments, making it difficult to implement effective internal controls.

then the claim that strong internal controls promote conservatism would become untenable. Instead, strong internal controls could provide no incentive or even a disincentive for firms to produce conservative accounting reports.

Further, regulators and standard-setters do not necessarily view conservatism as a desirable attribute of financial reporting. For instance, Barth (2008) points out that conservatism is not a qualitative characteristic of accounting information under the conceptual framework of the International Accounting Standards Board (IASB). The framework specifies that accounting information should be unbiased, and freedom from bias is an essential characteristic of reliability and of neutrality (IASB 2001, para. 31, 36). Conservatism, by allowing the deliberate understatement of assets or income and/or the deliberate overstatement of liabilities or expenses, renders financial statements not neutral and, therefore, jeopardizes the quality of reliability and unbiasedness desired by the conceptual framework.

The Financial Accounting Standards Board (FASB) asserts that the best interest of users is served by neutral reporting accompanied by appropriate disclosure of “the nature and extent of the uncertainty surrounding events and transactions reported to stockholders and others” (see SFAC No. 2, FASB 1980, para. 93; FASB 2000, para. 96, 97).⁸ Consistent with this perspective, the FASB is shifting its focus toward fair value accounting, lessening the asymmetric treatment of bad and good news in financial statements. If the objective of financial reporting is to provide information useful for making economic decisions (IASB 2001, para. 12), then having strong internal controls that ensure that the financial reports are unbiased would satisfy investors’ demand for neutral information. Hence, within this investor value-relevance framework, the argument that strong internal controls would promote conservative financial reports (which are essentially biased) seems tenuous.

Given these potentially conflicting perspectives, we empirically test the relation between internal controls and conservatism and test the null hypothesis that internal control quality is unrelated to conservatism. In order to further ascertain the nature of any relation between internal controls and conservatism, we perform intertemporal tests, as in Ashbaugh-Skaife et al. (2008), to determine whether firms with weak internal controls that subsequently show an improvement in internal controls exhibit different levels of conservatism from firms that fail to do so. This would help strengthen our finding on the relation between internal controls and conservatism. The null hypotheses we test are as follows:

H1: There is no relation between internal control quality and conservatism.

H2: Firms with weak internal controls that subsequently show an improvement in internal controls (i.e., remediate their internal control weaknesses) exhibit no difference in conservatism from firms that fail to do so.

III. RESEARCH DESIGN

Measures of Accounting Conservatism

Persistence of Earnings Changes (Basu 1997)

Our first measure to capture the differential timeliness of loss versus gain recognition is the persistence of earnings changes in Basu (1997). Basu (1997) shows that, relative to good news

⁸ For example, in SFAC No. 2 the FASB states, “Conservatism in financial reporting should no longer connote deliberate, consistent understatement of net assets and profits” (FASB 1980, para. 93). The FASB also states in the same paragraph that conservatism “became deeply ingrained and is still in evidence despite efforts over the past 40 years to change it.” In the joint exposure draft by the FASB and IASB, *Conceptual Framework for Financial Reporting* (FASB 2008, para. BC2.21), it states that “the boards concluded that describing *prudence* or *conservatism* as a qualitative characteristic or a desirable response to uncertainty would conflict with the quality of neutrality because ... an admonition to be prudent is likely to lead to a bias in the reported financial position and financial performance ... Accordingly, the proposed framework does not include *prudence* or *conservatism* as desirable qualities of financial reporting information” (emphasis in the original).

periods, conservatism results in lower persistence of earnings in bad news periods. The deferred recognition of relatively good news results in positive changes in income being less likely to reverse than negative earnings changes. This is because, from a time-series perspective, the bad news reflected in current earnings will appear as a transitory shock in the earnings process, whereas the effects of a current positive shock will be spread over the earnings of several future periods as anticipated gains are realized. The following model from Basu (1997) is used to estimate this relation:

$$\Delta NI_t = \alpha_0 + \alpha_1 D\Delta NI_{t-1} + \alpha_2 \Delta NI_{t-1} + \alpha_3 D\Delta NI_{t-1} * \Delta NI_{t-1} + \varepsilon \quad (1)$$

where firm i subscripts are omitted. ΔNI_t (ΔNI_{t-1}) is the change in net income before extraordinary items for firm i in fiscal year t ($t-1$) deflated by beginning-of-year total assets, and $D\Delta NI_{t-1}$ is an indicator variable that equals 1 if ΔNI_{t-1} is less than 0, and 0 otherwise.

The negative coefficient on $D\Delta NI_{t-1} * \Delta NI_{t-1}$ is consistent with timely loss recognition. We further develop model (1) into model (2) to test H1, where MW is an indicator variable that equals 1 if the firm has at least one MW in internal controls, and 0 otherwise. If MW firms report less (more) conservatively than firms without such weaknesses, then the coefficient on $D\Delta NI_{t-1} * \Delta NI_{t-1} * MW$ will be positive (negative); that is, MW firms have a reduced (increased) tendency to reverse negative earnings changes in the following period.

$$\begin{aligned} \Delta NI_t = & \alpha_0 + \alpha_1 D\Delta NI_{t-1} + \alpha_2 \Delta NI_{t-1} + \alpha_3 D\Delta NI_{t-1} * \Delta NI_{t-1} + \alpha_4 MW + \alpha_5 D\Delta NI_{t-1} * MW \\ & + \alpha_6 \Delta NI_{t-1} * MW + \alpha_7 D\Delta NI_{t-1} * \Delta NI_{t-1} * MW + \alpha_8 MB_t + \alpha_9 D\Delta NI_{t-1} * MB_t \\ & + \alpha_{10} \Delta NI_{t-1} * MB_t + \alpha_{11} D\Delta NI_{t-1} * \Delta NI_{t-1} * MB_t + \alpha_{12} LEV_t + \alpha_{13} D\Delta NI_{t-1} * LEV_t \\ & + \alpha_{14} \Delta NI_{t-1} * LEV_t + \alpha_{15} D\Delta NI_{t-1} * \Delta NI_{t-1} * LEV_t + \alpha_{16} SIZE_t + \alpha_{17} D\Delta NI_{t-1} * SIZE_t \\ & + \alpha_{18} \Delta NI_{t-1} * SIZE_t + \alpha_{19} D\Delta NI_{t-1} * \Delta NI_{t-1} * SIZE_t + \alpha_{20} LIT_t + \alpha_{21} D\Delta NI_{t-1} * LIT_t \\ & + \alpha_{22} \Delta NI_{t-1} * LIT_t + \alpha_{23} D\Delta NI_{t-1} * \Delta NI_{t-1} * LIT_t + \alpha_{24} IND_j + \alpha_{25} D\Delta NI_{t-1} * IND_j \\ & + \alpha_{26} \Delta NI_{t-1} * IND_j + \alpha_{27} D\Delta NI_{t-1} * \Delta NI_{t-1} * IND_j + \varepsilon \end{aligned} \quad (2)$$

where firm i subscripts are omitted.

Model (2) controls for firm characteristics that are used in prior research to proxy for the demand for conservatism. We include MB , the market value of equity divided by the book value of equity at the end of the fiscal year t , to control for the demand of conservatism arising from information asymmetry associated with a firm's growth option (LaFond and Watts 2008), and because studies document a negative association between conditional and unconditional conservatism (Givoly et al. 2007; Roychowdhury and Watts 2007). Firms with high levels of leverage tend to have greater bondholder and shareholder conflicts that, in turn, increase the contractual demand for conservatism (Ahmed et al. 2002; Zhang 2008). Hence, we include LEV , the sum of long-term debt and current liabilities deflated by total assets at the end of the fiscal year t .

LaFond and Watts (2008) argue that large firms produce more public information and have less information asymmetry, reducing the demand for conservative accounting. Hence, we control for firm size ($SIZE$) using the natural log of total assets at the end of the fiscal year t .⁹ We also control for litigation risk because it can increase managers' incentive to recognize losses in a more timely manner than gains (Basu 1997; Watts 2003). Following Francis et al. (1994), we define an

⁹ We mean-adjust all the continuous control variables in our regression models (i.e., MB , LEV , and $SIZE$) to reduce the problems with multicollinearity among the interaction terms (Neter et al. 1989; Aiken and West 1991). This procedure helps to make the magnitude of the coefficients more comparable to prior work.

indicator variable *LIT* that equals 1 if a firm operates in a litigious industry, and 0 otherwise.¹⁰ Finally, Givoly et al. (2007) contend that the degree of conservatism varies across different industries. Hence, we include industry indicators (*IND_j*) following the industry classification in Frankel et al. (2002).¹¹

H2 tests whether MW firms that remediate their MWs will exhibit different levels of conservatism from MW firms that continue to have these weaknesses. Following Ashbaugh-Skaife et al. (2008), we deem an unqualified SOX 404 opinion by the external auditor to objectively and unambiguously show that the firm has fully remediated its MWs. We test H2 using the sample of MW firms and replace *MW* with *FIXED* in model (2), where *FIXED* is an indicator variable that equals 1 if an MW firm receives an unqualified second SOX 404 report, and 0 otherwise. A negative (positive) coefficient on $\Delta NI_{t-1} * \Delta NI_{t-1} * FIXED$ indicates that MW firms that remediate their weaknesses exhibit greater (lower) conservatism than MW firms that still have weaknesses.

Accrual-Based Loss Recognition (Ball and Shivakumar 2005)

Ball and Shivakumar (2005) develop a model to describe the differential timeliness of gain and loss recognition that relies on the correlation between accruals and contemporaneous cash flows.¹² In their regression, as shown in model (3) below, the authors predict a positive coefficient on *DCFO * CFO* for accounting conservatism.

$$ACCRUAL = \gamma_0 + \gamma_1 DCFO + \gamma_2 CFO + \gamma_3 DCFO * CFO \quad (3)$$

where firm *i* and time *t* subscripts are omitted. *ACCRUAL* is net income before extraordinary items minus operating cash flows for firm *i* in fiscal year *t* deflated by beginning-of-year total assets. *CFO* is operating cash flows for firm *i* in fiscal year *t* deflated by beginning-of-year total assets. *DCFO* is an indicator variable that equals 1 if *CFO* is less than 0, and 0 otherwise. To examine H1, we develop model (4) as shown below. If MWs in internal controls result in lower (greater) accounting conservatism, then the accruals can less (more) effectively and timely reflect the future expectation of the negative change of cash flows; the coefficient on *DCFO * CFO * MW* is predicted to be negative (positive).

$$\begin{aligned} ACCRUAL = & \gamma_0 + \gamma_1 DCFO + \gamma_2 CFO + \gamma_3 DCFO * CFO + \gamma_4 MW + \gamma_5 DCFO * MW \\ & + \gamma_6 CFO * MW + \gamma_7 DCFO * CFO * MW + \gamma_8 MB + \gamma_9 DCFO * MB \\ & + \gamma_{10} CFO * MB + \gamma_{11} DCFO * CFO * MB + \gamma_{12} LEV + \gamma_{13} DCFO * LEV \\ & + \gamma_{14} CFO * LEV + \gamma_{15} DCFO * CFO * LEV + \gamma_{16} SIZE + \gamma_{17} DCFO * SIZE \end{aligned}$$

¹⁰ Consistent with Francis et al. (1994), firms with primary SIC codes of 2833–2836 (biotechnology), 3570–3577 (computer equipment), 3600–3674 (electronics), 5200–5961 (retailing), and 7370–7374 (computer services) are considered to be operating in a litigious industry.

¹¹ Following Frankel et al. (2002), industry membership is determined by the following SIC codes: agriculture (AGR, 0100–0999), mining and construction (MIN, 1000–1999, excluding 1300–1399), food (FOO, 2000–2111), textiles and printing/publishing (TEX, 2200–2799), chemicals (CHE, 2800–2824, 2840–2899), pharmaceuticals (PHA, 2830–2836), extractive (EXT, 2900–2999, 1300–1399), durable manufacturers (DUR, 3000–3999, excluding 3570–3579 and 3670–3679), transportation (TRA, 4000–4899), utilities (UTI, 4900–4999), retail (RET, 5000–5999), financial (FIN, 6000–6999), services (SER, 7000–8999, excluding 7370–7379) and computers (COM, 3570–3579, 3670–3679, 7370–7379).

¹² The authors assert a positive but asymmetric correlation between accruals and contemporaneous cash flows. This positive correlation arises because cash flow revisions in the current period tend to be positively correlated with the current revisions for expected future cash flows. Timely recognition of unrealized gains and losses is based on expected, not realized, cash flows, and is therefore accomplished through accruals. Consequently, timely gain and loss recognition will produce a positive correlation between accruals and current period cash flows. This correlation is asymmetric because losses, under conservative reporting, are likely to be recognized on a timelier basis than gains.

$$\begin{aligned}
& + \gamma_{18}CFO * SIZE + \gamma_{19}DCFO * CFO * SIZE + \gamma_{20}LIT + \gamma_{21}DCFO * LIT \\
& + \gamma_{22}CFO * LIT + \gamma_{23}DCFO * CFO * LIT + \gamma_{24}IND_j + \gamma_{25}DCFO * IND_j \\
& + \gamma_{26}CFO * IND_j + \gamma_{27}DCFO * CFO * IND_j + \varepsilon
\end{aligned} \quad (4)$$

where firm i and time t subscripts are omitted and all variables are as previously defined. We replace *MW* in model (4) with *FIXED* to test H2. If the remediation of MWs is associated with greater (lower) conservatism, then the coefficient on *DCFO * CFO * FIXED* is predicted to be positive (negative).

Timeliness of Earnings to News (Basu 1997)

Our third measure of conservatism is a firm's timeliness of earnings to news (Basu 1997). As shown in model (5) below, the timeliness of earnings is inferred from the responsiveness of accounting income to the change in market values. Negative (positive) market-adjusted stock returns are used as proxies for bad (good) news. The asymmetric recognition of economic losses relative to gains is captured by a positive coefficient on *DR * R*.

$$NI = \beta_0 + \beta_1 DR + \beta_2 R + \beta_3 DR * R + \varepsilon \quad (5)$$

where firm i and time t subscripts are omitted. *NI* is the net income before extraordinary items for firm i in fiscal year t , deflated by the beginning-of-year market value, R is the market-adjusted stock return for firm i over the fiscal year t , and *DR* is an indicator variable that equals 1 if R is less than 0, and 0 otherwise.¹³

To test H1, we develop model (6). A negative (positive) coefficient on *DR * R * MW* indicates that MW firms exhibit lower (higher) conservatism than firms without such weaknesses; that is, the MW firms have lower (higher) incremental timeliness of earnings to bad news than to good news. To test H2, we replace *MW* with *FIXED* in model (6). A positive (negative) coefficient on *DR * R * FIXED* would suggest that MW firms that remediate their weaknesses exhibit greater (lower) conservatism than MW firms that fail to do so.

$$\begin{aligned}
NI = & \beta_0 + \beta_1 DR + \beta_2 R + \beta_3 DR * R + \beta_4 MW + \beta_5 DR * MW + \beta_6 R * MW + \beta_7 DR * R * MW \\
& + \beta_8 MB + \beta_9 DR * MB + \beta_{10} R * MB + \beta_{11} DR * R * MB + \beta_{12} LEV + \beta_{13} DR * LEV \\
& + \beta_{14} R * LEV + \beta_{15} DR * R * LEV + \beta_{16} SIZE + \beta_{17} DR * SIZE + \beta_{18} R * SIZE \\
& + \beta_{19} DR * R * SIZE + \beta_{20} LIT + \beta_{21} DR * LIT + \beta_{22} R * LIT + \beta_{23} DR * R * LIT + \beta_{24} IND_j \\
& + \beta_{25} DR * IND_j + \beta_{26} R * IND_j + \beta_{27} DR * R * IND_j + \varepsilon
\end{aligned} \quad (6)$$

where firm i and time t subscripts are omitted and all variables are as previously defined.

Givoly et al. (2007) point out the importance of controlling for the information or disclosure environment when using the Basu's (1997) timeliness measure. The control variables in our

¹³ Dietrich et al. (2007) argue that the interpretation of model (5) is valid only when the market is efficient; in particular, if market returns cause earnings, and not the reverse. The authors also show that partitioning a regression sample with one of the regressors (R) may lead to biased inferences. However, Ryan (2006) notes that two well-known empirical results, the low R^2 s observed in contemporaneous returns-earnings regressions (Collins et al. 1997; Ely and Waymire 1999; Francis and Schipper 1999) and a large literature showing that returns typically reflect information on a timelier basis than earnings, indicate that the concern mentioned by Dietrich et al. (2007) is likely to induce a very tiny bias in the estimation of conservatism. Ryan (2006) recommends measuring returns over the fiscal year to partially remove the impact of the annual earnings announcement over stock prices, which occurs approximately three months after closing. He also recommends the use of market-adjusted returns, defined as raw returns minus the value-weighted CRSP market return, to create the partitioning dummy variable *DR* in the Basu regression. Hence, to mitigate the concerns of Dietrich et al. (2007), we measure returns over the fiscal year and use market-adjusted returns instead of raw returns.

regressions partially address this issue.¹⁴ Nonetheless, to further mitigate this concern, we control for other factors that are likely to proxy for the amount or quality of information disclosed by a firm—the annual frequency of management forecasts, the number of days from management forecast to fiscal year end date, audit quality (proxied by audit firm size), analyst following, and analyst forecast accuracy. The interpretation of our results for Basu's (1997) timeliness measure (untabulated) is not affected by the inclusion of these variables.

Sample Selection

Using the sample firms in Doyle et al. (2007a), we identify 1,098 firms that, under either SOX 302 or SOX 404, disclose at least one MW from January 2003 to November 2005.¹⁵ We focus on firms that disclosed MWs because the reporting of MWs is mandatory, whereas the reporting of significant deficiencies and control deficiencies is not (Doyle et al. 2007a).¹⁶ We deem a firm to have weak internal controls (i.e., low internal control quality) so long as it has at least one MW.

Table 1, Panel A summarizes the sample selection procedure for the descriptive statistics and for testing H1. We choose the fiscal years 2000 and 2001 to test H1 because these years just precede the enactment of SOX and, hence, avoid any confounding effects due to SOX. Our assumption is that MWs exist within the firm even before their disclosures from January 2003 to November 2005 (Doyle et al. 2007a). Based on the sample firms from Doyle et al. (2007a), we identify 1,230 (6,835) firm-year observations for the MW firms (control firms with no MW disclosures from January 2003 to November 2005) that have available data in Compustat and CRSP for fiscal years 2000 and 2001. Then we remove 84 (434) outlier observations for all the continuous variables at the top and bottom 0.5 percent levels, leaving us with 1,146 (6,401) firm-year observations for the MW firms (control firms) sample.¹⁷ In total, we have 7,547 firm-year observations for the descriptive statistics and H1 tests.

Table 1, Panel B summarizes the sample selection procedure for testing H2, our remediation test. We focus on the MW sample because remediation is not applicable to the control sample. Based on an initial 1,098 MW firms, we remove 473 firms that have no second SOX 404 opinions; that is, we are not sure whether the MWs are remediated. These firms are either non-accelerated filers that are not required to file SOX 404 opinions (289 firms) or they have terminated registra-

¹⁴ First, Givoly et al. (2007) contend that the timeliness measure is understated for firms where the information environment is characterized by a smooth and frequent arrival of news relative to firms for which the news arrives less frequently. The information environment of larger firms is more likely to conform to the former characterization. In this aspect, controlling for firm size addresses this concern. Second, Givoly et al. (2007) caution that if a highly litigious environment prompts management to preempt bad news by disclosing it early, the Basu's timeliness measure will under-estimate the degree of conservatism. Our inclusion of *LIT* as control in the regression mitigates this concern. Third, Givoly et al. (2007) point out that Basu's (1997) measure of conservatism is likely to be negatively related to unconditional conservatism. Hence, we include market-to-book value for control of unconditional conservatism. Fourth, Givoly et al. (2007) highlight that the information environment can vary across industries, which will affect the timeliness measure. We address this concern through the inclusion of industry-dummy variables as controls.

¹⁵ We thank Jeffrey Doyle, Weili Ge, and Sarah McVay for sharing the data. The data can be found at: <http://faculty.washington.edu/geweili/ICdata.html>.

¹⁶ Although both MWs and significant deficiencies are deficiencies in the design or operation of internal controls, significant deficiencies are less severe and are not required to be publicly disclosed under SOX 302 (SEC 2004). Hence, the disclosure of significant deficiencies is clearly voluntary. On the other hand, under SOX 302, if management identifies an MW in their controls, then they are precluded from reporting that the controls are effective and must disclose the identified MW. Hence, the disclosure of MWs is effectively mandatory. According to Doyle et al. (2007b), there is some ambiguity regarding whether SOX 302 certifications require the public disclosure of MWs and whether some firms might interpret the MW disclosure requirement under SOX 302 as voluntary. The authors' conclusion, from reading most of the SEC guidance, is that most firms are treating the disclosure as mandatory.

¹⁷ Because we use the same sample for the descriptive statistics and for all of the H1 tests (based on the three different conservatism measures), an outlier observation is removed from all the H1 tests, rather than from just the test to which it applies.

TABLE 1
Sample Selection Procedure

Panel A: Sample Selection for H1

	Observations
MW Sample	
Observations with available Compustat and CRSP data for 2000 and 2001	1,230
Less: outlier observations for all continuous variables at the top and bottom 0.5% levels	(84)
Total MW firm observations	<u>1,146</u>
Control Sample	
Observations with available Compustat and CRSP data for 2000 and 2001	6,835
Less: outlier observations for all continuous variables at the top and bottom 0.5% levels	(434)
Total control firm observations	<u>6,401</u>
Total firm-year observations used in descriptive statistics and testing H1	<u>7,547</u>

Panel B: Sample Selection for H2

	Observations
Identified MW firms from 2003 to 2005	1,098
less: firms without SOX 404 auditor opinions due to non-accelerated filers status	(289)
less: firms without SOX 404 auditor opinions due to termination of securities registration	(184)
less: firms with missing data in Compustat and CRSP	(109)
Total firm-year observations used in testing H2	<u>516*</u>

* Among the 516 firms with a second SOX 404 opinion, 366 received an unqualified auditor report and 150 received an adverse report.

tion of their securities (184 firms). We also remove 109 firms with missing data in Compustat and CRSP. This procedure yields a final sample of 516 firms for the H2 tests, with 366 unqualified opinions and 150 adverse opinions.

IV. RESULTS

Descriptive Statistics

Table 2, Panel A presents the descriptive statistics for the sample used to test H1, which is based on the firm-year observations for the fiscal years 2000 and 2001. Consistent with Doyle et al. (2007b) and Ashbaugh-Skaife et al. (2007), the table shows that MW firms are smaller, financially weaker, and more leveraged than the control firms. The MW firms are also more likely to operate in a litigious industry, and have more negative change in net income in year $t-1$ and more negative accruals than the control firms. We do not observe any differences in the change in net income in year t , stock returns, cash flow from operations, stock return, and market-to-book ratio between the MW and control firms. Table 2, Panel B presents both Pearson and Spearman correlation statistics; we do not find any unusual correlations among the independent variables in our regressions that warrant concern.

Multivariate Regression Results

Regression Results for H1

Table 3 presents the regression results of H1 based on the persistence of earnings changes measure. We first run a baseline regression model without the *MW* and its related terms. This allows us to assess the incremental explanatory power of internal control quality on conservatism once the *MW* terms are included into the estimations. The baseline regression results, presented in Column 1, show that the coefficient on $D\Delta NI_{t-1} * \Delta NI_{t-1}$ is -0.368 ($p < 0.01$), confirming that financial reporting is conservative in general. Column 2 shows that, after including the *MW* and the related terms, the adjusted R^2 of the regression increases from 24.0 percent to 24.6 percent. This result suggests that internal control quality provides incremental explanatory power to conservatism in addition to factors affecting the demand for conservatism. Column 2 also reveals that the coefficient on $D\Delta NI_{t-1} * \Delta NI_{t-1} * MW$ is 0.217 ($p = 0.01$), which implies that the MW firms have a reduced tendency to reverse negative earnings changes in the following period (i.e., are less conservative) relative to the control firms. This result is consistent with a positive relation between internal control quality and conservatism.

We repeat the above analyses using the accrual-based loss recognition measure and present the results in Table 4. Column 1 shows that the coefficient on $DCFO * CFO$ is 0.669 ($p < 0.01$), indicating the presence of reporting conservatism. Column 2 shows that the addition of *MW* and its related terms to the baseline regression increases the adjusted R^2 from 16.7 percent to 17.0 percent, suggesting that internal control quality provides incremental explanatory power to a firm's level of conservatism. The coefficient on $DCFO * CFO * MW$ is -0.149 ($p = 0.03$), indicating that the MW firms accrue less unrealized losses in the cash-loss year; this result provides further support that the MW firms exhibit lower conservatism than the control firms.

Finally, Table 5 presents the results of the same analyses using the timeliness of earnings to news measure. Column 1 shows that the coefficient on $DR * R$ is 0.402 ($p < 0.01$), consistent with the presence of conservative reporting. When we add the *MW* and its related terms to the baseline regression, the adjusted R^2 of the model increases from 28.0 percent to 29.0 percent. Column 2 further reveals that the coefficient on $DR * R * MW$ is -0.072 ($p = 0.06$), which implies that the earnings of the MW firms reflect unexpected losses in a less timely manner than the control firms (i.e., the MW firms are less conservative). Again, this finding is consistent with the above results

TABLE 2
Descriptive Statistics and Correlation Analysis

Panel A: Distributional Properties of Variables				
	<u>Mean</u>	<u>Median</u>	<u>Std. Dev.</u>	
MW Firms (n = 1,146 observations in 2000 and 2001)				
ΔNI_t	-0.02	0.00	0.16	<u>Q1</u>
ΔNI_{t-1}	-0.04**	0.01	0.12	-0.06
<i>ACCRUAL</i>	-0.08**	-0.06*	0.14	-0.18
<i>CFO</i>	0.05	0.06**	0.16	-0.13
<i>NI</i>	-0.06*	0.02**	0.24	-0.01
<i>R</i>	0.18	0.03*	0.84	-0.07
<i>MB</i>	2.69	1.64	3.69	-0.37
<i>LEV</i>	0.19**	0.13**	0.20	0.94
<i>SIZE-ACTUAL</i>	1811.17***	274.01*	6882.37	0.01
<i>LIT</i>	0.38***	0.00***	0.49	87.76
Control Firms (n = 6,401 observations in 2000 and 2001)				
ΔNI_t	-0.02	0.00	0.14	0.03
ΔNI_{t-1}	-0.03**	0.00	0.36	-0.14
<i>ACCRUAL</i>	-0.07**	-0.05*	0.13	-0.11
<i>CFO</i>	0.05	0.07**	0.17	0.00
<i>NI</i>	-0.05*	0.03**	0.23	-0.05
<i>R</i>	0.19	0.06*	0.78	-0.31
<i>MB</i>	2.68	1.66	3.90	0.92
<i>LEV</i>	0.18**	0.12**	0.19	0.00
<i>SIZE-ACTUAL</i>	3073.94***	301.64*	10489.78	74.01
<i>LIT</i>	0.30***	0.00***	0.46	0.00
				1457.23
				1.00

Panel B: Pearson Correlations (top) and Spearman Correlations (bottom)

	<u>ΔNI_t</u>	<u>ΔNI_{t-1}</u>	<u><i>ACCRUAL</i></u>	<u><i>CFO</i></u>	<u><i>NI</i></u>	<u><i>R</i></u>	<u><i>MB</i></u>	<u><i>LEV</i></u>	<u><i>SIZE</i></u>	<u><i>LIT</i></u>
ΔNI_t	1.00	-0.18†	0.48†	0.31†	0.44†	0.32†	0.04†	0.04†	0.05†	-0.12†
ΔNI_{t-1}	-0.06†	1.00	0.18†	0.24†	0.25†	0.05†	0.03†	0.04†	0.08†	-0.06†

(continued on next page)

Panel B: Pearson Correlations (top) and Spearman Correlations (bottom)

	ΔNI_t	ΔNI_{t-1}	ACC _{REAL}	CFO	NI	R	MB	LEV	SIZE	LIT
ACC _{REAL}	0.31†	0.13†	1.00	-0.07†	0.40†	0.11†	-0.05†	0.04†	0.09†	-0.16†
CFO	0.30†	0.27†	-0.28†	1.00	0.48†	0.24†	-0.03†	0.10†	0.28†	-0.18†
NI	0.41†	0.24†	0.46†	0.43†	1.00	0.27†	0.08†	0.07†	0.24†	-0.19†
R	0.22†	0.03†	0.11†	0.13†	0.26†	1.00	0.23†	0.01	0.02	-0.07†
MB	0.17†	0.15†	0.00	0.17†	-0.05†	0.30†	1.00	-0.07†	0.03†	0.16†
LEV	0.04†	0.01	0.01	0.09†	0.19†	0.08†	-0.11†	1.00	0.33†	-0.27†
SIZE	0.06†	0.08†	0.07†	0.25†	0.23†	0.12†	0.14†	0.42†	1.00	-0.18†
LIT	-0.09†	-0.03†	-0.14†	-0.13†	-0.27†	-0.15†	0.17†	-0.30†	-0.19†	1.00

*, **, *** Denote significant differences between the MW sample and the control sample at the 10 percent, 5 percent, and 1 percent levels, respectively, based on two-tailed tests. Differences in means (medians) are assessed using a t-test (Wilcoxon rank sum test).

† Indicates significance at the 5 percent level, two-tailed.

Panel A shows the distribution of the MW firms and the control firms with no MWs.

Panel B shows the correlations among the key variables (identified in Panel A) used in the empirical analyses. Pearson and Spearman correlations are found, respectively, above and below the diagonal.

Variable Definitions:

ΔNI_t = change in net income before extraordinary items for firm i in fiscal year t deflated by beginning-of-year total assets;

ΔNI_{t-1} = change in net income before extraordinary items for firm i in fiscal year $t-1$ deflated by beginning-of-year total assets;

ACC_{REAL} = net income before extraordinary items minus operating cash flows (Data308) for firm i in fiscal year t deflated by beginning-of-year total assets;

CFO = operating cash flows at the end of the fiscal year deflated by beginning-of-year total assets;

NI = net income before extraordinary items (Data123) for firm i in fiscal year t deflated by the beginning-of-year market value of equity (Data199 * Data25);

R = market-adjusted stock return for firm i over the fiscal year t ;

MB = market value of equity (Data199 * Data25) divided by the book value of equity (Data60) at the end of the fiscal year t ;

LEV = firm's leverage, measured by the sum of long-term debt (Data9) and debt in current liabilities (Data34) deflated by market value of equity at the end of the fiscal year t ;

SIZE = natural log of the total assets at the end of the fiscal year t . We use SIZE as the control variable in the regressions but for the descriptive statistics, we report the untransformed value of total assets (SIZE-ACTUAL); and

LIT = indicator variable that equals 1 if a firm operates in a litigious industry, and 0 otherwise. Consistent with Francis et al. (1994), firms with primary SIC codes of 2833-2836 (biotechnology), 3570-3577 (computer equipment), 3600-3674 (electronics), 5200-5961 (retailing), and 7370-7374 (computer services) are considered to be operating in a litigious industry.



TABLE 3
Regression Results of H1 using the Persistence of Earnings Changes Measure

Independent Variables	Pred. Sign	Baseline Regression (1)		Full Regression (2)	
		Coeff.	p-value	Coeff.	p-value
Intercept	?	-0.042***	<0.01	-0.043***	<0.01
$D\Delta NI_{t-1}$	?	-0.026***	<0.01	-0.025***	<0.01
ΔNI_{t-1}	?	-0.145***	0.01	-0.099*	<0.01
$D\Delta NI_{t-1} * \Delta NI_{t-1}$	-	-0.368***	<0.01	-0.412***	<0.01
<i>MW</i>	?			0.007	0.26
$D\Delta NI_{t-1} * MW$	?			-0.009	0.36
$\Delta NI_{t-1} * MW$	?			-0.246***	<0.01
$D\Delta NI_{t-1} * \Delta NI_{t-1} * MW$	?			0.217***	0.01
MB_t	?	0.005***	<0.01	0.005***	<0.01
$D\Delta NI_{t-1} * MB_t$	?	-0.008***	<0.01	-0.008***	<0.01
$\Delta NI_{t-1} * MB_t$	?	-0.017***	<0.01	-0.016***	<0.01
$D\Delta NI_{t-1} * \Delta NI_{t-1} * MB_t$	?	0.024***	<0.01	0.023***	<0.01
LEV_t	?	0.086***	<0.01	0.085***	<0.01
$D\Delta NI_{t-1} * LEV_t$	?	0.006	0.47	0.007	0.39
$\Delta NI_{t-1} * LEV_t$	?	0.395***	<0.01	0.406***	<0.01
$D\Delta NI_{t-1} * \Delta NI_{t-1} * LEV_t$	-	-0.374***	<0.01	-0.385***	<0.01
$SIZE_t$	?	0.005***	<0.01	0.004***	<0.01
$D\Delta NI_{t-1} * SIZE_t$	?	0.002	0.31	0.002	0.27
$\Delta NI_{t-1} * SIZE_t$	?	-0.096***	<0.01	-0.092***	<0.01
$D\Delta NI_{t-1} \Delta NI_{t-1} * SIZE_t$	+	0.149***	<0.01	0.144***	<0.01
LIT_t	?	-0.010	0.14	-0.010	0.16
$D\Delta NI_{t-1} * LIT_t$	?	-0.045***	<0.01	-0.046***	<0.01
$\Delta NI_{t-1} * LIT_t$	?	-0.177**	0.02	-0.187***	0.01
$D\Delta NI_{t-1} * \Delta NI_{t-1} * LIT_t$	-	0.137	0.15	0.147	0.13
Industry Effects		Not reported		Not reported	
Adjusted R ²		24.0%		24.6%	
Number of observations		7,547		7,547	

*, **, *** Denote significance at the 10 percent, 5 percent, and 1 percent levels, respectively, based on two-tailed tests. This table reports the regression results of H1 using the persistence of earnings changes measure of conservatism in Basu (1997). Column 1 (Column 2) shows the results without (with) the *MW* and its related terms. The dependent variable is ΔNI_t , which is the change in net income before extraordinary items for firm *i* in fiscal year *t* deflated by beginning-of-year total assets. *MW* is an indicator variable that equals 1 if the firm has at least one *MW* in internal controls, and 0 otherwise. The other variables are as previously defined. A positive (negative) coefficient on $D\Delta NI_{t-1} * \Delta NI_{t-1} * MW$ indicates that *MW* firms exhibit lower (greater) conservatism than the control firms.

of a positive relation between internal control quality and conservatism.¹⁸

¹⁸ Prior research shows that conservatism is associated with governance factors such as managerial ownership (LaFond and Roychowdhury 2008) and board of director characteristics (Ahmed and Duellman 2007). Hence, we also examine whether internal control quality has an incremental effect on conservatism after controlling for these factors. We repeat the H1 tests in Tables 3–5 by further adding CEO share ownership, institutional ownership, the duality of the CEO and

TABLE 4
Regression Results of H1 Using the Accrual-Based Loss Recognition Measure

Independent Variables	Pred. Sign	Baseline Regression (1)		Full Regression (2)	
		Coeff.	p-value	Coeff.	p-value
Intercept	?	-0.020***	0.01	-0.014*	0.09
DCFO	?	0.002	0.86	-0.026*	0.10
CFO	-	-0.502***	<0.01	-0.534***	<0.01
DCFO * CFO	+	0.669***	<0.01	0.643***	<0.01
MW	?			-0.013*	0.07
DCFO * MW	?			-0.006	0.64
CFO * MW	?			0.048	0.34
DCFO * CFO * MW	?			-0.149**	0.03
MB	?	0.003***	<0.01	0.003***	<0.01
DCFO * MB	?	-0.004***	<0.01	-0.004***	0.01
CFO * MB	?	0.000	1.00	0.000	0.95
DCFO * CFO * MB	?	0.000	0.98	0.000	0.93
LEV	?	-0.023	0.11	-0.021	0.15
DCFO * LEV	?	0.004	0.88	-0.006	0.82
CFO * LEV	?	-0.232**	0.03	-0.241**	0.03
DCFO * CFO * LEV	+	0.455***	<0.01	0.438***	<0.01
SIZE	?	0.003**	0.03	0.003**	0.03
DCFO * SIZE	?	0.000	0.97	0.000	0.97
CFO * SIZE	?	0.006	0.56	0.005	0.59
DCFO * CFO * SIZE	-	-0.070***	<0.01	-0.070***	<0.01
LIT	?	-0.009	0.40	-0.008	0.46
DCFO * LIT	?	-0.024	0.16	-0.023	0.17
CFO * LIT	?	0.046	0.58	0.041	0.62
DCFO * CFO * LIT	+	0.083	0.42	0.104	0.31
Industry Effects		Not reported		Not reported	
Adjusted R ²		16.7%		17.0%	
Number of observations		7,547		7,547	

*, **, *** Denote significance at the 10 percent, 5 percent, and 1 percent levels, respectively, based on two-tailed tests.

This table reports the regression results of H1 using the accrual-based loss recognition measure of conservatism in Ball and Shivakumar (2005). Column 1 (Column 2) shows the results without (with) the *MW* and its related terms. The dependent variable is *ACCRUAL*, which is net income before extraordinary items minus operating cash flows for firm *i* in fiscal year *t* deflated by beginning-of-year total assets. The other variables are as previously defined. A negative (positive) coefficient on *DCFO * CFO * MW* indicates that *MW* firms exhibit lower (greater) conservatism than the control firms.

Our findings on the control variables are generally consistent with prior research. For instance, the coefficients on $D\Delta NI_{t-1} * \Delta NI_{t-1} * MB$ and $DR * R * MB$ confirm the negative relation

Chairman positions, board size, board independence, the share ownership of the independent directors, and the average outside directorships of the independent directors. The sample size for these tests is 1,920, which is much smaller than that used in the analyses in Tables 3–5. This is because we obtain the data on managerial ownership from the ExecuComp database and the governance characteristics from the RiskMetrics database. Both databases provide information on S&P 1500 companies only. Our inferences on the results for H1 (untabulated) remain unchanged. Specifically, the coefficients on $D\Delta NI_{t-1} * \Delta NI_{t-1} * MW$, $DCFO * CFO * MW$, and $DR * R * MW$ are 0.423 ($p = 0.06$), -0.510 ($p = 0.03$), and -0.089 ($p = 0.07$), respectively.

TABLE 5
Regression Results of H1 using the Timeliness of Earnings to News Measure

Independent Variables	Pred. Sign	Baseline Regression (1)		Full Regression (2)	
		Coeff.	p-value	Coeff.	p-value
Intercept	?	-0.011*	0.07	-0.009	0.13
<i>DR</i>	?	0.049***	<0.01	0.052***	<0.01
<i>R</i>	?	0.011**	0.05	0.013**	0.03
<i>DR * R</i>	+	0.402***	<0.01	0.414***	<0.01
<i>MW</i>	?			-0.010	0.39
<i>DR * MW</i>	?			-0.017	0.41
<i>R * MW</i>	?			-0.011	0.29
<i>DR * R * MW</i>	?			-0.072*	0.06
<i>MB</i>	?	0.001	0.37	0.001	0.37
<i>DR * MB</i>	?	-0.004**	0.05	-0.004*	0.06
<i>R * MB</i>	?	-0.001	0.23	-0.001	0.22
<i>DR * R * MB</i>	?	-0.020***	<0.01	-0.020***	<0.01
<i>LEV</i>	?	-0.037	0.12	-0.033	0.15
<i>DR * LEV</i>	?	0.018	0.67	0.017	0.69
<i>R * LEV</i>	?	0.032	0.19	0.030	0.22
<i>DR * R * LEV</i>	+	0.372***	<0.01	0.378***	<0.01
<i>SIZE</i>	?	0.010***	<0.01	0.009***	<0.01
<i>DR * SIZE</i>	?	-0.005	0.21	-0.004	0.23
<i>R * SIZE</i>	?	-0.001	0.84	-0.001	0.79
<i>DR * R * SIZE</i>	-	-0.077***	<0.01	-0.076***	<0.01
<i>LIT</i>	?	-0.063***	<0.01	-0.062***	<0.01
<i>DR * LIT</i>	?	0.011	0.57	0.012	0.55
<i>R * LIT</i>	?	0.013	0.22	0.014	0.19
<i>DR * R * LIT</i>	+	-0.034	0.32	-0.033	0.34
Industry Effects		Not reported		Not reported	
Adjusted R ²		28.0%		29.0%	
Number of observations		7,547		7,547	

*, **, *** Denote significance at the 10 percent, 5 percent, and 1 percent levels, respectively, based on two-tailed tests.

This table reports the regression results of H1 using the timeliness of earnings to news measure of conservatism in Basu (1997). Column 1 (Column 2) shows the results without (with) the *MW* and its related terms. The dependent variable is *NI*, which is the net income before extraordinary items for firm *i* in fiscal year *t* deflated by the beginning-of-year market value of equity. The other variables are as previously defined. A negative (positive) coefficient on *DR * R * MW* indicates that *MW* firms exhibit lower (greater) conservatism than the control firms.

between conservatism and unconditional conservatism documented in other studies.¹⁹ The coefficients on $D\Delta NI_{t-1} * \Delta NI_{t-1} * LEV$, $DCFO * CFO * LEV$, and $DR * R * LEV$ are also consistent with prior research that shows that more highly leveraged firms exhibit more conservatism. The

¹⁹ We also use another measure of unconditional conservatism devised by Penman and Zhang (2002), which is the amount of a firm's "hidden" reserves estimated using the sum of the value of the LIFO reserve, the research and development reserve, and the advertising reserve. Using this alternative measure, we continue to find a significant and positive relation between internal control quality and conservatism, and a significant and negative relation between unconditional conservatism and our conservatism measures.

coefficients on $D\Delta NI_{t-1} * \Delta NI_{t-1} * SIZE$, $DCFO * CFO * SIZE$, and $DR * R * SIZE$ suggest that larger firms have lower conservatism, consistent with the aggregation effect discussed in Givoly et al. (2007) and with the information asymmetry hypothesis discussed in LaFond and Watts (2008). Finally, similar to LaFond and Roychowdhury (2008), we do not find a significant relation between the level of litigation risks and conservatism.

Regression Results for H2

Table 6 reports the regression results for testing H2 using the persistence of earnings changes measure (Column 1), accrual-based loss recognition measure (Column 2), and timeliness of earnings to news measure (Column 3). For brevity, we present only the coefficients of the three-way interaction terms for the control variables. Column 1 reveals that the coefficient on $D\Delta NI_{t-1} * \Delta NI_{t-1} * FIXED$ is -0.622 ($p = 0.03$). This result shows that MW firms that remediate

TABLE 6
Regression Results of H2

	Pred. Sign	Persistence of Earnings Changes A = $D\Delta NI_{t-1}$ B = ΔNI_{t-1} (1)		Accrual-Based Loss Recognition A = $DCFO$ B = CFO (2)		Timeliness of Earnings to News A = DR B = R (3)	
		Coeff.	p-value	Coeff.	p-value	Coeff.	p-value
Intercept	?	-0.004	0.85	-0.020	0.28	-0.017	0.56
A	?	-0.004	0.91	-0.111***	<0.01	-0.003	0.95
B	?	-0.553***	<0.01	-0.803***	<0.01	0.046	0.39
A * B	-, +, +	0.160	0.54	0.436*	0.10	0.281***	<0.01
FIXED	?	0.014	0.59	0.016	0.40	0.039	0.19
A * FIXED	?	-0.018	0.63	0.097***	<0.01	0.069	0.11
B * FIXED	?	0.240	0.29	0.205	0.17	-0.002	0.97
A * B * FIXED	?	-0.622**	0.03	0.748***	<0.01	0.186*	0.07
A * B * MB	?	0.050	0.15	-0.041***	0.01	0.013	0.16
A * B * LEV	-, +, +	2.082**	0.05	0.397	0.38	0.391*	0.07
A * B * SIZE	+, -, -	0.021	0.85	-0.201***	0.01	0.036	0.24
A * B * LIT	-, +, +	-0.700**	0.05	-0.393	0.17	-0.241***	0.04
Industry Effects		Not reported		Not reported		Not reported	
Adjusted R ²		30.7%		43.1%		33.6%	
Number of observations		516		516		516	

*, **, *** Denote significance at the 10 percent, 5 percent, and 1 percent levels, respectively, based on two-tailed tests. This table reports the regression results of H2 using the persistence of earnings changes measure (Column 1), accrual-based loss recognition measure (Column 2), and timeliness of earnings to news measure (Column 3). The dependent variables for Columns 1, 2, and 3 are ΔNI , $ACCRUAL$, and NI , respectively. $FIXED$ is an indicator variable that equals 1 if an MW firm receives an unqualified second SOX 404 report, and 0 otherwise. All the other variables are as previously defined. For brevity, we present only the coefficients of the three-way interaction terms of the control variables. A negative (positive, positive) coefficient on $D\Delta NI_{t-1} * \Delta NI_{t-1} * FIXED$ ($DCFO * CFO * FIXED$, $DR * R * FIXED$) indicates that MW firms that remediate their weaknesses at the time of the second SOX 404 report exhibit greater conservatism than MW firms that fail to do so. A positive (negative, negative) coefficient on $D\Delta NI_{t-1} * \Delta NI_{t-1} * FIXED$ ($DCFO * CFO * FIXED$, $DR * R * FIXED$) indicates that MW firms that remediate their weaknesses at the time of the second SOX 404 report exhibit lower conservatism than MW firms that fail to do so.

their weaknesses at the time of the second SOX 404 report exhibit greater conservatism than MW firms that continue to have weaknesses at the time of the report. The results using the other two measures of conservatism are consistent with this finding. Specifically, the coefficient on $DCFO * CFO * FIXED$ in Column 2 is 0.748 ($p < 0.01$), and the coefficient on $DR * R * FIXED$ in Column 3 is 0.186 ($p = 0.07$). Hence, the regression results on H2 using all three measures of conservatism strengthen our earlier finding of a positive relation between internal control quality and conservatism.

In summary, the empirical findings using all three measures of conservatism show that internal control quality provides incremental explanatory power to a firm's level of conservatism. We find consistent evidence of a positive relation between internal control quality and conservatism. Specifically, firms with MWs exhibit lower conservatism than firms without such weaknesses, and MW firms that subsequently remediate their weaknesses exhibit greater conservatism than MW firms that fail to do so. Taken together, our results show that effective internal controls act as a mechanism that facilitates conservatism, consistent with the agency framework of positive accounting theory that recognizes conservatism as a governance mechanism (Watts 2003; Ball and Shivakumar 2005). In the next section, we describe additional analyses used to test the robustness of these results.

V. ADDITIONAL ANALYSES

Self-Selection of Internal Control Quality (H1)

Prior research suggests that MW firms are likely to be systematically different from other firms (Ashbaugh-Skaife et al. 2007; Doyle et al. 2007b). This could result in potential self-selection bias in our observed samples for our H1 tests. Although the tests on the change in the status of internal controls mitigate this concern, we further econometrically control for this bias using three methods. First, we use Heckman's (1979) two-stage procedure, in which the first stage comprises a probit regression of *MW* on its determinants. These determinants largely follow those of Doyle et al. (2007b) and Ashbaugh-Skaife et al. (2007), and are detailed in Table 7.²⁰ From this regression that identifies the likelihood of a firm being "selected" as an MW firm, we calculate the inverse Mills ratio, *LAMBDA* (see Heckman 1979; Leuz and Verrecchia 2000), and include it in our main regressions to control for the likelihood of self-selection into the MW group.

Second, we create a matched sample based on the predicted probabilities from the first-stage probit regression. This method, known as propensity score-matching (LaLonde 1986), creates a non-MW control sample with the same predicted probabilities of having an MW (i.e., it considers all predictive variables together to estimate the same propensity of having an MW). The propensity score-matching produces a matched sample of 850 control firm observations, resulting in a combined sample of 1,700 firm observations.²¹ Third, we include the determinants of MWs, as well as their two-way and three-way interactions with the parameters of the three conservatism measures, as additional control variables in the main regressions presented earlier. This method controls for the possibility that the factors that produce the MWs could also be driving observed differences in conservatism.

Table 7 presents the results of the above three tests for the persistence of earnings changes measure. For brevity, we present only the results of the three-way interaction terms for the control variables. Column 1 shows the second-stage regression results of the Heckman two-stage proce-

²⁰ The first-stage regression results (untabulated) show that financial health (*LOSS*), rapid growth (*EXTR_SALES*), audit quality (*BIGN*), the occurrence of a restatement (*RESTATE*), and the occurrence of an auditor change (*AUDCHANGE*) are significantly associated with the presence of MWs. Details of the results are available from the authors upon request.

²¹ The search algorithm finds the propensity scores closest to our MW firm observations until the first decimal place. The algorithm is available upon request.

TABLE 7
Regression Results of H1 Using the Persistence of Earnings Changes Measure and after Controlling for Self-Selection of Internal Control Quality

	Pred. Sign	Heckman Second-Stage Regression (1)		Propensity Score-Matching (2)		Controlling for Determinants of MWs (3)	
		Coeff.	p-value	Coeff.	p-value	Coeff.	p-value
Intercept	?	-0.086***	<0.01	-0.047***	<0.01	-0.033***	0.01
$D\Delta NI_{t-1}$	?	-0.022***	0.01	-0.038*	0.08	-0.037*	0.06
ΔNI_{t-1}	?	-0.014	0.82	-0.286***	<0.01	0.128	0.37
$D\Delta NI_{t-1} * \Delta NI_{t-1}$	-	-0.718***	<0.01	0.143	0.36	-0.384*	0.08
<i>MW</i>	?	0.016**	0.02	0.014	0.25	0.011*	0.10
$D\Delta NI_{t-1} * MW$	?	-0.016	0.15	0.003	0.86	-0.008	0.47
$\Delta NI_{t-1} * MW$	?	-0.422***	<0.01	-0.141*	0.06	-0.346***	<0.01
$D\Delta NI_{t-1} * \Delta NI_{t-1} * MW$	?	0.503***	<0.01	0.276***	0.01	0.445***	<0.01
$D\Delta NI_{t-1} * \Delta NI_{t-1} * MB_t$	?	0.023***	0.01	-0.024**	0.02	0.012	0.17
$D\Delta NI_{t-1} * \Delta NI_{t-1} * LEV_t$	-	-0.553***	<0.01	-0.839***	<0.01	-0.444***	<0.01
$D\Delta NI_{t-1} * \Delta NI_{t-1} * SIZE_t$	+	0.045**	0.05	0.208***	<0.01	0.014	0.59
$D\Delta NI_{t-1} * \Delta NI_{t-1} * LIT_t$	-	0.154	0.16	-0.761***	<0.01	0.088	0.37
<i>LAMBDA</i>	?	0.025***	<0.01				
$D\Delta NI_{t-1} * \Delta NI_{t-1} * AGE_t$	?					0.035	0.47
$D\Delta NI_{t-1} * \Delta NI_{t-1} * LOSS_t$	?					-0.485***	<0.01
$D\Delta NI_{t-1} * \Delta NI_{t-1} * SEGMENT_t$	?					-0.026	0.21
$D\Delta NI_{t-1} * \Delta NI_{t-1} * FOREIGN_t$	?					-0.223***	0.01
$D\Delta NI_{t-1} * \Delta NI_{t-1} * MA_t$	?					0.549*	0.10
$D\Delta NI_{t-1} * \Delta NI_{t-1} * RESTR_t$	?					2.079***	<0.01
$D\Delta NI_{t-1} * \Delta NI_{t-1} * EXTR\_SALES_t$	?					0.283***	<0.01
$D\Delta NI_{t-1} * \Delta NI_{t-1} * BIGN_t$	?					0.121	0.30
$D\Delta NI_{t-1} * \Delta NI_{t-1} * RESTATE_t$	?					0.824***	<0.01
$D\Delta NI_{t-1} * \Delta NI_{t-1} * AUDCHANGE_t$	?					-0.143*	0.08

(continued on next page)

TABLE 7 (continued)

	Heckman Second-Stage Regression (1)		Propensity Score-Matching (2)		Controlling for Determinants of MWs (3)		
	Pred. Sign	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value
Industry Effects		Not reported		Not reported		Not reported	
Adjusted R ²		27.3%		17.2%		32.7%	
Number of observations		5,439		1,700		5,439	

*, **, *** Denote significance at the 10 percent, 5 percent, and 1 percent levels, respectively, based on two-tailed tests.

This table reports the results of the regressions that control for self-selection of internal control quality using the persistence of earnings changes measure. The dependent variable is ΔMI_t , which is the change in net income before extraordinary items for firm i in fiscal year t deflated by beginning-of-year total assets. For brevity, other than MW_t , we present only the coefficients of the three-way interaction terms of the control variables. Column 1 presents the second stage results of the Heckman (1979) two-stage procedure, after controlling for $LAMBDA$. $LAMBDA$ is the inverse Mill ratio calculated from the first-stage regression in which we regress MW on its determinants. Column 2 presents the results of the propensity-score-matching approach. Column 3 presents the results after controlling for the determinants of MW s. We use the following determinants of MW s. Firm size is measured by the log of market capitalization, which is share price (Data25) multiplied by number of shares outstanding (Data199). Because we have already controlled for $SIZE$, we do not add market capitalization as further control in Column 3. The other variables are defined as follows or as previously defined.

Variable Definitions:

AGE = firm age, the number of years since the firm appears in CRSP database;

$LOSS$ = indicator variable that equals 1 if net income before extraordinary items (Data123) in years t and $t-1$ sum to less than 0, and 0 otherwise;

$SEGMENT$ = log of the sum of the number of operating and geographic segments reported by the Compustat Segments database for the firm in year t ;

$FOREIGN$ = indicator variable that equals 1 if the firm has a non-zero foreign currency translation (Data150), and 0 otherwise;

MA = indicator variable that equals 1 if the firm has a non-zero merger and acquisition activity (Data360);

$RESTR$ = aggregate restructuring charges [Data376 $\times$ (-1)] in years t and $t-1$, scaled by the firm's market capitalization at the end of the fiscal year;

$EXTR\_SALES$ = indicator variable that equals 1 if year-over-year industry-adjusted sales growth (Data12) falls into the top quintile, and 0 otherwise;

$BIGN$ = indicator variable that equals 1 if the firm engaged one of the largest four audit firms, and 0 otherwise. Largest four audit firms include PWC, Deloitte & Touche, Ernst & Young, and KPMG;

$RESTATE$ = indicator variable that equals 1 if the firm had a restatement in the 12-month period before the disclosure of MW s, and 0 otherwise; and

$AUDCHANGE$ = indicator variable that equals 1 if the firm changed auditor during the 12-month period before the disclosure of MW s, and 0 otherwise.

*, **, *** Denote significance at the 10 percent, 5 percent, and 1 percent levels, respectively, based on two-tailed tests.

This table reports the results of the regressions that control for self-selection of internal control quality using the persistence of earnings changes measure. The dependent variable is ΔNI_t , which is the change in net income before extraordinary items for firm i in fiscal year t deflated by beginning-of-year total assets. For brevity, other than MW , we present only the coefficients of the three-way interaction terms of the control variables. Column 1 presents the second stage results of the Heckman (1979) two-stage procedure, after controlling for $LAMBDA$. $LAMBDA$ is the inverse Mill ratio calculated from the first-stage regression in which we regress MW on its determinants. Column 2 presents the results of the propensity-score-matching approach. Column 3 presents the results after controlling for the determinants of MWs. We use the following determinants of MWs. Firm size is measured by the log of market capitalization, which is share price (Data25) multiplied by number of shares outstanding (Data199). Because we have already controlled for $SIZE$, we do not add market capitalization as further control in Column 3. The other variables are defined as follows or as previously defined.

Variable Definitions:

AGE = firm age, the number of years since the firm appears in CRSP database;

$LOSS$ = indicator variable that equals 1 if net income before extraordinary items (Data123) in years t and $t-1$ sum to less than 0, and 0 otherwise;

$SEGMENT$ = log of the sum of the number of operating and geographic segments reported by the Compustat Segments database for the firm in year t ;

$FOREIGN$ = indicator variable that equals 1 if the firm has a non-zero foreign currency translation (Data150), and 0 otherwise;

MA = indicator variable that equals 1 if the firm has a non-zero merger and acquisition activity (Data360);

$RESTR$ = aggregate restructuring charges [Data376 $\times (-1)$] in years t and $t-1$, scaled by the firm's market capitalization at the end of the fiscal year;

$EXTR\_SALES$ = indicator variable that equals 1 if year-over-year industry-adjusted sales growth (Data12) falls into the top quintile, and 0 otherwise;

$BIGN$ = indicator variable that equals 1 if the firm engaged one of the largest four audit firms, and 0 otherwise. Largest four audit firms include PWC, Deloitte & Touche, Ernst & Young, and KPMG;

$RESTATE$ = indicator variable that equals 1 if the firm had a restatement in the 12-month period before the disclosure of MWs, and 0 otherwise; and

$AUDCHANGE$ = indicator variable that equals 1 if the firm changed auditor during the 12-month period before the disclosure of MWs, and 0 otherwise.

ture after controlling for *LAMBDA*. Column 2 shows the propensity score-matching results and Column 3 shows the results after controlling for the determinants of MWs. In all three regressions, we find that the coefficient on $D\Delta NI_{t-1} * \Delta NI_{t-1} * MW$ remains positive and significant, confirming our earlier findings in Table 3. When we repeat these analyses for the other two measures of conservatism, the results for all the three tests (untabulated) show that the coefficients on $DCFO * CFO * MW$ and $DR * R * MW$ remain negative and significant ($p < 0.10$).²² Taken together, we interpret these findings as suggesting that self-selection issues are not likely to drive the results of H1 that we obtained earlier.

Self-Selection of Improvement in Internal Control Quality (H2)

It is also possible that MW firms that choose to improve their internal controls can introduce a possible self-selection bias into our tests of H2. To control for this potential bias, we use the Heckman two-stage procedure and estimate a probit regression of *FIXED* on the same determinants of MWs as above, reasoning that the characteristics that lead to the incidence of MWs are also likely to cause difficulties for these firms to remediate MWs. We also include governance characteristics because firms with better governance are more likely to remediate MWs (Goh 2009).²³ We hand-collect data on the board and audit committee characteristics from the proxy statements of the 516 MW firms used earlier to test H2, and are left with 437 firms with available data.

Table 8 reports the results after including the inverse Mills ratio, *LAMBDA*, in our previous regressions used to test H2. For brevity, we omit the results for the control variables. The results in Table 8 are similar to those when we do not control for self-selection. Specifically, the coefficients on $D\Delta NI_{t-1} * \Delta NI_{t-1} * FIXED$, $DCFO * CFO * FIXED$, and $DR * R * FIXED$ remain significant and have the same signs as those in Table 6. We continue to find similar significant results (untabulated) when we include the determinants of *FIXED*, instead of *LAMBDA* as additional control variables in the regressions. In sum, we conclude that the results for H2 are unlikely to be driven by self-selection in the improvement in internal control quality.

Alternative Measures of Conditional Conservatism

Khan and Watts (2009) suggest a firm-specific measure of conservatism (*C_SCORE*) that is based on Basu's (1997) measure of asymmetric timeliness of earnings. We follow their approach and derive the *C_SCORE* measure, with higher *C_SCORE* indicating greater conservatism. The regression model and results are presented in Table 9.²⁴ In summary, the results using the *C_SCORE* measure confirm our earlier findings. Specifically, the coefficient on *MW* in Column 1

²² To save space, we do not tabulate the results. However, these results, as well as all other results that we do not tabulate, are available from the authors upon request.

²³ These governance characteristics include board size, board independence, board meeting frequency, audit committee size, audit committee accounting financial expertise, audit committee meeting frequency, the duality of the CEO and Chairman positions, board commitment (proxied by the number of outside directorships in public companies held by the board directors), and board diligence (proxied by the proportion of directors who fail to attend more than 75 percent of the board meetings). The first-stage regression results show that firm profitability, the occurrence of a restatement, board independence, and audit committee meeting frequency are significantly associated with the likelihood of the remediation of MWs. Details on the definitions of the governance characteristics and the first-stage regression results are available from the authors upon request.

²⁴ Khan and Watts (2009) show that firms with high uncertainty and long investment cycles have future gains that are less verifiable and are more susceptible to gaming, generating a higher contracting demand for conservatism. Hence, we control for information uncertainty using *STD* (defined as the standard deviation of monthly stock returns) and control for the length of the investment cycle using *CYCLE* (defined as the depreciation expenses deflated by lagged asset). We control for *AGE* (firm age in years) because younger firms tend to have larger information asymmetry that is associated with more conservatism. Finally, we include *SPREAD* (the bid-ask spread scaled by the midpoint of the spread) to proxy for information asymmetry that is positively related to conservatism (LaFond and Watts 2008).

TABLE 8
Regression Results of H2 after Controlling for Self-Selection of Remediation of MWs
Using the Heckman Two-Stage Procedure

	Pred. Sign	Persistence of Earnings Changes A = $D\Delta NI_{t-1}$ B = ΔNI_{t-1} (1)		Accrual-Based Loss Recognition A = $DCFO$ B = CFO (2)		Timeliness of Earnings to News A = DR B = R (3)	
		Coeff.	p-value	Coeff.	p-value	Coeff.	p-value
Intercept	?	0.004	0.88	-0.010	0.65	0.014	0.89
A	?	0.000	0.99	-0.102***	<0.01	0.100	0.41
B	?	0.522***	<0.01	-0.496***	<0.01	0.132	0.45
A * B	-, +, +	0.230	0.37	0.457*	0.09	0.455	0.11
FIXED	?	0.009	0.73	0.005	0.78	0.049	0.63
A * FIXED	?	-0.001	0.98	0.088***	<0.01	0.166	0.24
B * FIXED	?	0.348	0.13	0.055	0.69	-0.031	0.87
A * B * FIXED	?	-0.905**	0.01	0.664***	<0.01	0.893**	0.01
LAMBDA	?	-0.021	0.35	-0.042***	<0.01	-0.144**	0.05
Control Variables		Not reported		Not reported		Not reported	
Industry Effects		Not reported		Not reported		Not reported	
Adjusted R ²		24.6%		49.4%		23.1%	
Number of observations		437		437		437	

*, **, *** Denote significance at the 10 percent, 5 percent, and 1 percent levels, respectively, based on two-tailed tests.

This table reports the results of the regressions that control for self-selection in the remediation of MWs. The results are from the second-stage regression of the Heckman (1979) two-stage procedure, after controlling for *LAMBDA*. *LAMBDA* is the inverse Mill ratio calculated from the first-stage regression in which we regress *FIXED* on the determinants of MW remediation. The determinants of MW remediation include the determinants of MWs and the governance characteristics detailed in footnote 23. For brevity, we do not report the results for the control variables. Columns 1, 2, and 3 present the results using the persistence of earnings changes measure, accrual-based loss recognition measure, and the timeliness of earnings to news measure, respectively. In Column 1, the dependent variable is ΔNI_t , which is the change in net income before extraordinary items for firm *i* in fiscal year *t* deflated by beginning-of-year total assets. In Column 2, the dependent variable is *ACCRUAL*, which is net income before extraordinary items minus operating cash flows for firm *i* in fiscal year *t* deflated by beginning-of-year total assets. In Column 3, the dependent variable is *NI*, which is the net income before extraordinary items for firm *i* in fiscal year *t* deflated by the beginning-of-year market value of equity. All the other variables are as previously defined.

is -0.007 (*p* = 0.02), consistent with MW firms exhibiting lower levels of conservatism than the control firms. The coefficient on *FIXED* in Column 2 is 0.288 (*p* = 0.03), consistent with MW firms that remediate their weaknesses exhibiting greater conservatism than MW firms that continue to have weaknesses.

To further corroborate our findings, we use a second accrual-based measure of conservatism (*CON-ACC*) that is suggested by Givoly and Hayn (2000). This measure is computed as income before extraordinary items less cash flows from operations plus depreciation expense deflated by average total assets, averaged over a three-year period before year *t*.²⁵ The lower the *CON-ACC*,

²⁵ The intuition behind this measure is that conservative accounting results in persistently negative accruals. Averaging the accruals over a number of periods ensures that the effects of any temporary large accruals are mitigated, as accruals tend to reverse within a one- to three-year period (Richardson et al. 2005). For the H1 test, we measure *CON-ACC* over the

TABLE 9
Regression Results Using the Firm-Specific Conservatism Measure in
Khan and Watts (2009)

H1:

$$C\_SCORE = \lambda_0 + \lambda_1 MW + \lambda_2 SIZE + \lambda_3 LEV + \lambda_4 MB + \lambda_5 STD + \lambda_6 CYCLE + \lambda_7 AGE + \lambda_8 SPREAD + \varepsilon.$$

H2:

$$C\_SCORE = \lambda_0 + \lambda_1 FIXED + \lambda_2 SIZE + \lambda_3 LEV + \lambda_4 MB + \lambda_5 STD + \lambda_6 CYCLE + \lambda_7 AGE + \lambda_8 SPREAD + \varepsilon.$$

	Pred. Sign	Test of H1 (1)		Test of H2 (2)	
		Coeff.	p-value	Coeff.	p-value
Intercept	?	0.484***	<0.01	1.408*	0.07
MW	?	-0.007**	0.02		
FIXED	?			0.288**	0.03
SIZE	-	-0.067***	<0.01	-0.184	0.14
LEV	+	0.233***	<0.01	1.847***	<0.01
MB	?	-0.001**	0.02	-0.025**	0.03
STD	+	0.063***	<0.01	1.373	0.26
CYCLE	-	-0.340***	<0.01	-6.569***	<0.01
AGE	-	0.000	0.64	-0.008	0.19
SPREAD	+	-0.011**	0.02	2.086***	0.01
Industry Effects		Not reported		Not reported	
Adjusted R2		73.8%		7.0%	
Number of observations		5836		500	

*, **, *** Denote significance at the 10 percent, 5 percent, and 1 percent levels, respectively, based on two-tailed tests. This table reports the results of the regressions that examine H1 and H2 using the firm-specific conservatism measure *C_SCORE* in Khan and Watts (2009). The dependent variable is *C_SCORE*, with higher *C_SCORE* indicating greater conservatism. The other variables are defined as follows or as previously defined.

Variable Definitions:

- STD* = standard deviation of monthly stock returns;
CYCLE = depreciation expenses deflated by the beginning-of-year total assets, which is a decreasing measure of the length of the investment cycle;
AGE = firm age, the number of years since the firm appears in CRSP database; and
SPREAD = bid-ask spread scaled by the midpoint of the spread, obtained from CRSP.

three-year period before each 2000 and 2001 firm-year observation. For the H2 test, we measure *CON-ACC* over the three-year period right after the year of the second SOX 404 report. In these tests, we control for firm size (*SIZE*), leverage (*LEV*), and litigation risks (*LIT*) as before. Consistent with Ahmed et al. (2002), we control for growth in sales (*SALESGROWTH*) because it will affect accruals such as inventory and receivables, which in turn affect *CON-ACC*. We control for research and development plus advertising expenditures (*RDADV*) because they are likely to capture economic rents generated by assets-in-place, growth opportunities, and GAAP mandated conservatism (Ahmed 1994). Finally, we control for profitability (*CFO*) as proxied by the operating cash flows deflated by beginning-of-year total assets, because Ahmed et al. (2002) argue that profitable firms tend to use more conservative accounting.

the more conservative the firm's accounting. Table 10 presents the results using *CON-ACC*, which is consistent with all our earlier findings. Specifically, the coefficient on *MW* in Column 1 is 0.010 ($p = 0.05$), implying that *MW* firms have larger operating accruals (i.e., lower conservatism) than the control firms. Further, the coefficient on *FIXED* in Column 2 is -0.005 ($p = 0.08$), which indicates that *MW* firms that remediate their weaknesses exhibit greater conservatism than *MW* firms that fail to do so. In sum, our earlier findings on H1 and H2 are robust to alternative conservatism measures suggested by Khan and Watts (2009) and Givoly and Hayn (2000).

Expanded Sample for H2 Tests

In our H2 tests in Table 6, we remove 473 firms without SOX 404 opinions; that is, firms that are either non-accelerated filers (289 firms) or have terminated the registration of their securities

TABLE 10
Regression Results Using the Accrual-Based Conservatism Measure in
Givoly and Hayn (2000)

H1:

$$CON-ACC = \theta_0 + \theta_1 MW + \theta_2 SIZE + \theta_3 SALES GROWTH + \theta_4 RDADV + \theta_5 CFO + \theta_6 LEV + \theta_7 LIT + \varepsilon.$$

H2:

$$CON-ACC = \theta_0 + \theta_1 FIXED + \theta_2 SIZE + \theta_3 SALES GROWTH + \theta_4 RDADV + \theta_5 CFO + \theta_6 LEV + \theta_7 LIT + \varepsilon.$$

	Pred. Sign	Test of H1 (1)		Test of H2 (2)	
		Coeff.	p-value	Coeff.	p-value
Intercept	?	-0.079***	<0.01	-0.021***	<0.01
<i>MW</i>	?	0.010**	0.05		
<i>FIXED</i>	?			-0.005*	0.08
<i>SIZE</i>	+	0.007***	<0.01	0.002**	0.03
<i>SALES GROWTH</i>	+	-0.003	0.38	0.005*	0.10
<i>RDADV</i>	+	-0.012***	<0.01	0.000	0.63
<i>CFO</i>	-	0.068***	<0.01	0.052***	<0.01
<i>LEV</i>	-	-0.009	0.45	-0.010	0.18
<i>LIT</i>	-	-0.029***	<0.01	-0.008**	0.04
Industry Effects		Not reported		Not reported	
Adjusted R2		3.5%		10.4%	
Number of observations		7,634		407	

*, **, *** Denote significance at the 10 percent, 5 percent, and 1 percent levels, respectively, based on two-tailed tests. This table reports the results of the regressions that examine H1 and H2 using the accrual-based conservatism measure in Givoly and Hayn (2000). The dependent variable is *CON-ACC*, which is income before extraordinary items less cash flows from operations plus depreciation expense deflated by average total assets, and averaged over a three-year period. The lower the *CON-ACC*, the greater the conservatism. The other variables are defined as follows or as previously defined.

Variable Definitions:

SALES GROWTH = percentage of annual growth in total sales (Data12); and

RDADV = research and development expenditures plus advertising expense divided by sales.

(184 firms). This procedure results in a loss of roughly half of our *MWs* observations. To recover some of these lost observations, we relax our requirements to determine the remediation of *MWs* by using the SOX 302 opinions for non-accelerated filers. For these non-accelerated filers, we rely on the SOX 302 opinions at the time of the second SOX 404 report that would have been issued if the firm had been an accelerated filer. This procedure yields larger sample sizes of 814, 800, and 651 firms with available data from Compustat and/or CRSP for the persistence of earnings changes measure, accruals-based loss recognition measure, and timeliness of earnings to news measure, respectively.²⁶ We then replicate the H2 tests in Table 6 using these expanded samples, and find that the inferences of our earlier results remain unchanged. Specifically, untabulated results show that the coefficients on $D\Delta NI_{t-1} * \Delta NI_{t-1} * FIXED$, $DCFO * CFO * FIXED$, and $DR * R * FIXED$ are -1.393 ($p < 0.01$), 0.557 ($p = 0.06$), and 0.607 ($p = 0.01$), respectively.

Contextual Analysis of the Nature of *MWs*

Finally, to provide a stronger link between internal controls and conservatism, we further explore the nature of the *MWs*. Using the arguments found in Ball (2001), we expect firms with *MWs* pertaining to investment and/or asset management related decisions to have a greater incentive to avoid timely loss recognition.²⁷ In addition, following the arguments in Doyle et al. (2007a), we expect company-level *MWs* to be more severe and pervasive than other types of *MWs*, and consequently to result in even lower conservatism.²⁸ Hence, we code firms with at least one of the above types of *MWs* as “severe *MW* firms” and test whether these firms exhibit lower conservatism than the other *MW* firms. To conduct the analyses, we replace *MW* with *SEVERE* in the regression models (2), (4), and (6), where *SEVERE* is an indicator variable that equals 1 for severe *MW* firms, and 0 otherwise. Using the *MW* firms observations from fiscal years 2001 and 2002, the regression results (untabulated) reveal that the coefficients on $D\Delta NI_{t-1} * \Delta NI_{t-1} * SEVERE$, $DCFO * CFO * SEVERE$, and $DR * R * SEVERE$ are 0.677 ($p < 0.01$), -0.069 ($p = 0.57$), and -0.127 ($p = 0.07$), respectively. Hence, the contextual analysis of the nature of *MWs* shows some evidence that severe *MW* firms exhibit lower conservatism than the other *MW* firms. This evidence provides further empirical support of a positive relation between internal control quality and conservatism.

²⁶ The significantly smaller sample size for the timeliness of earnings to news measure is due to the additional requirements for returns data from CRSP. Unlike before, we do not require that all three tests using different measures of conservatism have the same sample size. In other words, we do not impose the same data limitations to all measures of conservatism. This allows us to obtain the maximum sample size for each measure of conservatism to conduct the additional analyses on H2.

²⁷ Ball (2001) states that both debt contracting and corporate governance give rise to the demand for accounting conservatism, because conservatism can incorporate into financial statements adverse information about future cash flows of investment that is observable only to managers. *MWs* in the investment and/or asset management-related decisions can impede this governance role of conservatism. For instance, control weaknesses over the computation of the future cash flows of acquired investments and the lack of staff expertise in estimating future cash flows for asset impairment tests make it harder for the accounting system to incorporate into financial statements adverse information about future cash flows, which will lead to lower level of conservatism. The examples of *MWs* found in the Appendix are illustrative of these types of *MWs*.

²⁸ Following Doyle et al. (2007a), company-level *MWs* include override by senior management and ineffective control environment. These *MWs* are more severe and pervasive than account-specific *MWs*, which include inadequate internal controls for accounting for loss contingencies, deficiencies in the documentation of a receivables securitization program, and inadequate internal controls over the application of new accounting principles or the application of existing accounting principles to new transactions. Consistent with Doyle et al. (2007a), we also consider firms that have at least three account-specific *MWs* to have company-level *MWs*. We would like to thank Jeffrey Doyle, Weili Ge, and Sarah McVay for sharing the data on the classification of *MWs*.

VI. CONCLUSION

Our results show that firms that disclose material weaknesses (MWs) exhibit lower conditional conservatism than firms without such weaknesses, and MW firms that remediate their weaknesses exhibit greater conservatism than MW firms that fail to do so. Collectively, our results are consistent with strong internal controls facilitating conservatism. These findings extend the literature on the reporting effects of strong versus weak internal controls. In addition, our study contributes to the broader literature that tests the extent to which conservatism is shaped by a firm's contracting and governance environment.

One limitation of our study is that for our tests of H2, we remove nearly half of the 1,098 MW firms that do not have SOX 404 opinions due to their non-accelerated filer status and due to termination of securities registration. It is unclear how the addition of these firms would affect the results of H2. Another limitation is that our findings cannot directly shed light on the current debate regarding the costs and benefits of SOX. Our study informs only of the value of internal control quality, but not of the effectiveness of the internal control reporting requirements under SOX.

APPENDIX

EXAMPLES OF MATERIAL WEAKNESSES THAT COULD IMPEDE THE TIMELY RECOGNITION OF LOSSES

NEOMAGIC CORP (28 April, 2005)

NeoMagic Corporation has a material weakness as of January 31, 2005 related to its process of evaluating long-lived assets for impairment, including insufficient controls over the review and documentation of key assumptions used in preparing forecasted cash flows used to support NeoMagic Corporation's impairment analyses. This material weakness resulted in an audit adjustment that was recorded by NeoMagic Corporation to recognize the impairment of intangible assets.

OPENTV CORP (16 March, 2005)

We did not have sufficient internal personnel and technical expertise to properly apply accounting principles to certain non-routine matters; in particular, we applied certain provisions of Statement of Financial Accounting Standards No. 142, *Goodwill and Other Intangibles* (SFAS No. 142), in an incorrect manner when conducting our annual analysis for potential impairment of goodwill, which led to an error that was identified and subsequently corrected.

CLEARONE COMMUNICATIONS INC (18 August, 2005)

We have a material weakness related to the tracking and valuation of inventory, including controls to identify and properly account for obsolete inventory. Our accounting policies and practices over tracking and valuation of inventory, including controls to identify and properly account for slow-moving, obsolete inventory were inconsistent with GAAP. This material weakness resulted in errors in recording inventories at the lower of cost or market, and errors for inventory shrinkage.

We have a material weakness in accounting for non-routine transactions, which include business combinations, discontinued operations, sale of a business unit (other than discontinued operations), and evaluation and recognition of impairment charges. This material weakness resulted in improper purchase price allocations in business combinations, improper amortization and depreciation of long-lived assets, improper identification and recording of activities related to discontinued operations, improper recording and reporting the sale of business units, improper evaluation of triggering events associated with impairment of long-lived assets (including annual impairment tests for goodwill); and improper calculating and recording of impairment charges.

ASCONI CORP (27 May 2004)

We lacked controls to ensure that agreements, contracts, and other documents relating to investments in subsidiaries or investees were provided to and reviewed by accounting and financial reporting personnel on a timely basis to ensure that the financial reporting and disclosure implications of such transactions could be considered and reflected in the financial statements in the proper periods.

FINANCIAL INDUSTRIES CORP (10 June, 2003)

Management identified a significant number of internal control weaknesses in several key areas that had resulted in material misstatement of financial results. These include weaknesses in the following areas, among others: deferred policy acquisition costs; present value of future profits of acquired businesses; investment accounting; consolidation process; purchase accounting; and intercompany accounting.

REFERENCES

- Ahmed, A. S. 1994. Accounting earnings and future economic rents: An empirical analysis. *Journal of Accounting and Economics* 17 (3): 377–400.
- , B. K. Billings, R. M. Morton, and M. Stanford-Harris. 2002. The role of accounting conservatism in mitigating bondholder-shareholder conflicts over dividend policy and in reducing debt costs. *The Accounting Review* 77 (4): 867–890.
- , and S. Duellman. 2007. Accounting conservatism and board of director characteristics: An empirical analysis. *Journal of Accounting and Economics* 43 (2–3): 411–437.
- Aiken, L., and S. West. 1991. *Multiple Regression: Testing and Interpreting Interactions*. Thousand Oaks, CA: SAGE Publications Inc.
- Armstrong, C. S., W. R. Guay, and J. P. Weber. 2010. The role of information and financial reporting in corporate governance and debt contracting. Working paper, University of Pennsylvania.
- Ashbaugh-Skaife, H., D. Collins, and W. Kinney. 2007. The discovery and reporting of internal control deficiencies prior to SOX-mandated audits. *Journal of Accounting and Economics* 44 (1–2): 166–192.
- , ———, and R. LaFond. 2008. The effect of SOX internal control deficiencies and their remediation on accruals quality. *The Accounting Review* 83 (1): 217–250.
- , ———, ———, and ———. 2009. The effect of SOX internal control deficiencies on firm risk and cost of equity. *Journal of Accounting Research* 47 (1): 1–43.
- Ball, R., 2001. Infrastructure requirements for an economically efficient system of public financial reporting and disclosure. *Brookings-Wharton Papers on Financial Services*: 127–169.
- , S. P. Kothari, and A. Robin. 2000. The effect of international institutional factors on properties of accounting earnings. *Journal of Accounting and Economics* 29 (1): 1–51.
- , A. Robin, and J. Wu. 2003. Incentives versus standards: Properties of accounting income in four East Asian countries. *Journal of Accounting and Economics* 36 (1–3): 235–270.
- , and L. Shivakumar. 2005. Earnings quality in UK private firms: Comparative loss recognition timeliness. *Journal of Accounting and Economics* 39 (1): 83–128.
- , and ———. 2006. The role of accruals in asymmetrically timely gain and loss recognition. *Journal of Accounting Research* 44 (2): 207–242.
- Barth, M. 2008. Global financial reporting: Implications for U.S. academics. *The Accounting Review* 83 (5): 1159–1179.
- Basu, S. 1997. The conservatism principle and the asymmetric timeliness of earnings. *Journal of Accounting and Economics* 24 (1): 3–37.
- , L. Hwang, and C. Jan. 2001. Differences in conservatism between Big Eight and non-Big Eight auditors. Working paper, Emory University, Chinese University of Hong Kong, and California State University, Hayward.
- Beaver, W. H., and S. G. Ryan. 2005. Conditional and unconditional conservatism: Concepts and modeling.

- Review of Accounting Studies* 10 (2–3): 269–309.
- Beneish, M. D., M. Billings, and L. Hodder. 2008. Internal control weaknesses and information uncertainty. *The Accounting Review* 83 (3): 665–703.
- Bushman, R., and A. Smith. 2001. Financial accounting information and corporate governance. *Journal of Accounting and Economics* 32 (1–3): 237–333.
- , and J. Piotroski. 2006. Financial reporting incentives for conservative accounting: The influence of legal and political institutions. *Journal of Accounting and Economics* 42 (1–2): 107–148.
- Collins, D., E. Maydew, and I. Weiss. 1997. Changes in the value-relevance of earnings and book values over the past forty years. *Journal of Accounting and Economics* 24 (1): 39–67.
- Committee of Sponsoring Organizations of the Treadway Commission (COSO). 1992. *Internal Control—Integrated Framework*. The framework. New York, NY: COSO.
- Dietrich, J., K. Muller, and E. Riedl. 2007. Asymmetric timeliness tests of accounting conservatism. *Review of Accounting Studies* 12 (1): 95–124.
- Doyle, J., W. Ge, and S. McVay. 2007a. Accruals quality and internal control over financial reporting. *The Accounting Review* 82 (5): 1141–1170.
- , ———, and ———. 2007b. Determinants of weaknesses in internal control over financial reporting. *Journal of Accounting and Economics* 44 (1–2): 193–223.
- Ely, K., and G. Waymire. 1999. Accounting standard-setting organizations and earnings relevance: Longitudinal evidence from NYSE common stocks, 1927–1973. *Journal of Accounting Research* 37 (2): 293–317.
- Fama, E. F., and M. C. Jensen. 1983. Separation of ownership and control. *The Journal of Law & Economics* 26 (2): 301–325.
- Feng, M., C. Li, and S. McVay. 2009. Internal control and management guidance. *Journal of Accounting and Economics* 48 (2–3): 190–209.
- Financial Accounting Standards Board (FASB). 1980. *Qualitative Characteristics of Accounting Information*. Statement of Financial Accounting Concepts No. 2. Norwalk, CT: FASB.
- . 2000. *Using Cash Flow Information and Present Value in Accounting Measurements*. Statement of Financial Accounting Concepts No. 7. Norwalk, CT: FASB.
- . 2008. *Exposure Draft, Conceptual Framework for Financial Reporting: The Objective of Financial Reporting and Qualitative Characteristics and Constraints of Decision-Useful Financial Reporting Information*. May 28. Norwalk, CT: FASB.
- Francis, J. R., D. R. Philbrick, and K. Schipper. 1994. Shareholder litigation and corporate disclosures. *Journal of Accounting Research* 32 (2): 137–164.
- , and K. Schipper. 1999. Have financial statements lost their relevance? *Journal of Accounting Research* 37 (2): 319–352.
- Frankel, R. M., M. F. Johnson, and K. K. Nelson. 2002. The relation between auditors' fees for non-audit services and earnings management. *The Accounting Review* 77 (Supplement): 71–106.
- Givoly, D., and C. Hayn. 2000. The changing time-series properties of earnings, cash flows and accruals: Has financial reporting become more conservative? *Journal of Accounting and Economics* 29 (3): 287–320.
- , ———, and A. Natarajan. 2007. Measuring reporting conservatism. *The Accounting Review* 82 (1): 65–106.
- , ———, and S. P. Katz. 2010. Does public ownership of equity improve earnings quality? *The Accounting Review* 85 (1): 195–225.
- Goh, B. W. 2009. Audit committees, boards of directors, and remediation of material weaknesses in internal control. *Contemporary Accounting Research* 26 (2): 549–579.
- Guay, W., and R. Verrecchia. 2006. Discussion of: The role of accruals in asymmetrically timely gain and loss recognition. *Journal of Accounting Research* 44 (2): 243–255.
- Heckman, J. J. 1979. Sample selection bias as a specification error. *Econometrica* 47 (1): 153–162.
- Holthausen, R. W., and R. L. Watts. 2001. The relevance of value-relevance literature for financial accounting standard-setting. *Journal of Accounting and Economics* 31 (1–3): 3–75.
- International Accounting Standards Board (IASB). 2001. *Framework for the Preparation and Presentation of Financial Statements*. London, UK: IASB.

- Jensen, M. C. 1993. The modern industrial revolution, exit and the failure of internal control systems. *The Journal of Finance* 48 (3): 831–880.
- Khan, M., and R. L. Watts. 2009. Estimation and empirical properties of a firm-year measure of conservatism. *Journal of Accounting and Economics* 48 (2–3): 132–150.
- LaFond, R., and S. Roychowdhury. 2008. Managerial ownership and accounting conservatism. *Journal of Accounting Research* 46 (1): 101–135.
- , and R. L. Watts. 2008. The information role of conservative financial statements. *The Accounting Review* 83 (2): 447–478.
- LaLonde, R. 1986. Evaluating the econometric evaluations of training programs with experimental data. *The American Economic Review* 76 (4): 604–620.
- Leuz, C., and R. Verrecchia. 2000. The economic consequences of increased disclosure. *Journal of Accounting Research* 38 (Supplement): 91–124.
- Levitt, A. 1998. The numbers game. Remarks delivered at the NYU Center for Law and Business, New York, NY, September 28.
- Neter, J., W. Wasserman, and M. H. Kutner. 1989. *Applied Linear Regression Models*. 2nd edition. Homewood, IL: Irwin.
- Nikolaev, V. 2010. Debt covenants and accounting conservatism. *Journal of Accounting Research* 48 (1): 51–89.
- Ogneva, M., K. R. Subramanyam, and K. Raghunandan. 2007. Internal control weaknesses and cost of equity: Evidence from SOX Section 404 disclosures. *The Accounting Review* 82 (5): 1255–1297.
- Penman, S., and X. Zhang. 2002. Accounting conservatism, the quality of earnings and stock returns. *The Accounting Review* 77 (2): 237–264.
- Public Company Accounting Oversight Board (PCAOB). 2004. *An Audit of Internal Control Over Financial Reporting Performed in Conjunction with an Audit of Financial Statements*. Auditing Standards No. 2. Washington, D.C.: PCAOB.
- Richardson, S. A., R. G. Sloan, M. T. Soliman, and I. A. Tuna. 2005. Accrual reliability, earnings persistence and stock prices. *Journal of Accounting and Economics* 39 (3): 437–485.
- Roychowdhury, S., and R. Watts. 2007. Asymmetric timeliness of earnings, market-to-book and conservatism in financial reporting. *Journal of Accounting and Economics* 44 (1–2): 2–31.
- Ryan, S. G. 2006. Identifying conditional conservatism. *European Accounting Review* 15 (4): 511–525.
- Securities and Exchange Commission (SEC). 2002. *Final Rule: Certification of Disclosure in Companies' Quarterly and Annual Reports*. Release Nos. 33-8124; 34-46427. Washington, D.C.: SEC
- . 2003. *Final Rule: Management's Reports on Internal Control over Financial Reporting and Certification of Disclosure in Exchange Act Periodic Reports*. Release Nos. 33-8238; 34-47986. Washington D.C.: SEC August 14.
- . 2004. *Management's Reports on Internal Control over Financial Reporting and Certification of Disclosure in Exchange Act Periodic Reports-Frequently Asked Questions*. Washington D.C.: SEC.
- Watts, R. L. 2003. Conservatism in accounting part 1: Explanations and implications. *Accounting Horizons* 17 (3): 207–221.
- , and J. L. Zimmerman. 1986. *Positive Accounting Theory*. Englewood Cliffs, NJ: Prentice Hall Inc.
- Zhang, J. 2008. The contracting benefits of accounting conservatism to lenders and borrowers. *Journal of Accounting and Economics* 45 (1): 27–54.



Relative Performance Evaluation and Related Peer Groups in Executive Compensation Contracts

Guojin Gong

The Pennsylvania State University

Laura Yue Li

University of Illinois at Urbana–Champaign

Jae Yong Shin

Seoul National University

ABSTRACT: This study examines the explicit use of relative performance evaluation (RPE) in executive compensation contracts and the selection of RPE peers. Using S&P 1500 firms' first proxy disclosures under the SEC's 2006 executive compensation disclosure rules, we find that about 25 percent of our sample firms explicitly use RPE in setting executive compensation. We demonstrate that a lack of knowledge of both actual peer-group composition and the link between RPE-based performance targets and future peer performance significantly hinder the traditional Implicit test from detecting RPE use. We also find that firms consider both costs and benefits of RPE as an incentive mechanism when deciding to use RPE. Finally, both efficient contracting and rent extraction considerations influence RPE peer selection, with the relative importance of these competing considerations depending on RPE firms' performance.

Keywords: *executive compensation; relative performance evaluation; peer group; SEC regulation.*

Data Availability: *Data are available from public sources indicated in the text.*

We thank Ana Albuquerque, Mary Ellen Carter, Ingolf Dittmann, Ron Dye, Zhaoyang Gu, Ranjani Krishnan (editor), Ronald Masulis, two anonymous referees, and workshop participants at the 6th Accounting Research Workshop at the Universities of Fribourg and Bern, the 2009 Chinese Accounting Professors' Association of North America Meeting, Northwestern University, The Pennsylvania State University, University of Illinois at Urbana–Champaign, University of Kentucky, University of Wisconsin–Madison, and Vanderbilt University for their helpful comments. We also thank Jean Liang, Maria Revelo, Tony Shin, Jeff Wang, Steve Wu, and Victor Yang for their valuable research assistance. Professor Shin acknowledges financial support from the Institute of Management Research at Seoul National University.

Editor's note: Accepted by Ranjani Krishnan.

*Submitted: March 2009
Accepted: October 2010
Published Online: April 2011*

I. INTRODUCTION

Relative performance evaluation (RPE) entails the use of peer performance in evaluating an agent's performance. Based upon the premise that peer performance captures common exogenous shocks, the major theoretical benefit of RPE is to insulate an agent from common risk, while providing the principal with a more informative measure to assess the agent's actions. Consequently, RPE is expected to lead to risk-sharing benefits in incentive contracting (e.g., Lazear and Rosen 1981; Holmstrom 1982; Nalebuff and Stiglitz 1983).

Despite RPE's theoretical appeal, prior empirical research offers mixed evidence on the use of RPE in executive compensation contracts based on an *implicit* approach.¹ The implicit approach infers RPE use by regressing executive pay on industry performance across a population of firms, and thus relies on simplified assumptions concerning RPE contract details (such as RPE peer-group composition, performance metrics used in RPE, and components of pay covered by RPE). These assumptions unavoidably introduce measurement errors into implicit tests (Murphy 1999; Bannister and Newman 2003). More importantly, studies that rely on the implicit approach are unable to examine the execution of RPE contracts (such as peer selection), which is essential to evaluating the costs and benefits of RPE as an incentive mechanism (Matsumura and Shin 2006; Bol et al. 2010).

The SEC's 2006 executive compensation disclosure rules require firms to provide details on how performance targets (including *relative* performance targets) are used in setting executive pay. We exploit this mandatory disclosure regime to examine the explicit use of RPE and related peer groups. We first examine the extent to which simplified assumptions concerning RPE contract details affect inferences drawn from the traditional implicit approach. We then examine firms' decisions to explicitly use RPE in setting executive pay. Finally, we focus on RPE peer groups, an essential element of RPE contracts, to investigate factors that might influence the selection of peers for RPE purposes.

Using S&P 1500 firms' first proxy disclosures under the SEC's new disclosure rules, we find that about 25 percent of S&P 1500 firms explicitly use RPE in setting executive compensation. When using the implicit approach (such as in Albuquerque [2009]), where RPE peers are matched on both industry and size, we do not find evidence of RPE use in S&P 1500 firms in our 2006 sample.² We further show that this implicit approach is unable to detect RPE use even among firms that claim to use RPE in setting executive pay. However, after incorporating disclosed peer-group composition, we find a significantly negative association between CEO pay and stock performance for disclosed peers, supporting the use of RPE. Moreover, we find that failure to account for the link between RPE-based performance targets and future peer performance hinders the detection of RPE use in implicit tests. These results illustrate that simplified assumptions about RPE contract details cloud inferences drawn from implicit tests.

Turning to firms' decisions to explicitly use RPE, we find that firms exposed to higher common risk, operating in less-concentrated industries, having fewer growth opportunities, and hiring less-wealthy CEOs are more likely to use RPE. These findings are in line with economic theories predicting that a firm's use of RPE increases with the firm's exposure to common risk and decreases with the CEO's self-hedging ability (e.g., Lazear and Rosen 1981; Holmstrom 1982; Garvey and Milbourn 2003). Moreover, firms that are larger, have more independent and larger boards, and hire compensation consultants are more likely to use RPE. These results reveal the

¹ Examples include Antle and Smith (1986), Barro and Barro (1990), Gibbons and Murphy (1990), Janakiraman et al. (1992), Joh (1999), Aggarwal and Samwick (1999), Garvey and Milbourn (2003), Rajgopal et al. (2006), and Albuquerque (2007, 2009).

² In untabulated results, using the implicit approach as per Albuquerque (2009), we replicate Albuquerque's main finding and obtain a significantly negative coefficient on industry-size-matched peer returns over the 1992–2005 period.

importance of board structure and compensation consultants in facilitating the use of RPE. The overall evidence supports the view that firms consider both costs and benefits of RPE as an incentive mechanism when deciding to use RPE.

Next, among firms using RPE in the pay-setting process, we examine whether firms choose RPE peers in a manner consistent with efficient contracting or rent extraction. Under the efficient-contracting perspective, the selection of RPE peers should consider potential peers' ability to capture common risk (e.g., Holmstrom 1982). Tournament theory further suggests selecting RPE peers such that RPE firms' and peer firms' executives have equal chances to "win," given similar effort levels (e.g., Lazear and Rosen 1981). Under the rent-extraction perspective, executives have incentives to choose RPE peers that they expect to underperform (e.g., Gibbons and Murphy 1990; Murphy 2001), as choosing peers with poor expected performance can inflate RPE firms' peer-adjusted performance and, hence, increase executive compensation.

The evidence supports both efficient contracting and rent-extraction behavior in the peer-selection process. Consistent with efficient contracting, RPE firms tend to choose peers with higher common-risk-reduction ability and with greater similarity. These findings are generally more pronounced for RPE firms that outperform their industry than for underperforming RPE firms. At the same time, we find that RPE firms are more likely to select peers that are expected to perform poorly (based on analysts' target price forecasts). Interestingly, this self-serving bias in peer selection is more significant among underperforming RPE firms. Taken together, our findings suggest that RPE firms' superior performance leads to efficient-contracting considerations, whereas RPE firms' underperformance leads to rent-seeking behavior in the peer-selection process. Executives in underperforming RPE firms likely have stronger incentives to behave opportunistically in the peer-selection process because they are more concerned about potential reputation loss and pay reduction as a result of poor performance. When we examine intertemporal changes in peer-group composition, we find some evidence that added peers are more similar in size to RPE firms, but exhibit poorer performance than dropped peers.

Our study contributes to the executive compensation literature in three ways. First, our evidence facilitates a better understanding of the "RPE puzzle" that has long plagued researchers (Murphy 1999; Prendergast 1999). Our findings suggest that the RPE puzzle of mixed empirical support for RPE use arises at least in part from simplified assumptions underlying the traditional implicit approach. We demonstrate that incorporating RPE contract details (such as peer-group composition and the link between RPE-based performance targets and future peer performance) can significantly alter inferences drawn from implicit tests. Our evidence thus cautions against relying on the traditional implicit approach to draw definitive conclusions regarding RPE use.

Second, we provide the first large-sample U.S. evidence on the explicit use of RPE. We find that about 25 percent of our sample firms use RPE in setting executive compensation. Our findings on firms' decisions to use RPE confirm related economic theories, while also revealing the importance of board structure and compensation consultants in facilitating RPE use. This stands in contrast to Carter et al. (2009), who do not find evidence in support of related economic theories when studying explicit RPE use in performance-vested equity plans for a sample of U.K. firms.

Third, we provide the first large-sample evidence on the selection of RPE peers. We document that both efficient contracting (captured by common risk and economic similarity) and rent extraction (measured by expected peer performance) explain the likelihood of being selected as a RPE peer. These findings shed light on the execution of RPE contracts, which is essential to understanding the overall costs and benefits of RPE as an incentive device.

Our evidence also adds to the growing literature on the use of peer groups in executive compensation contracts. Existing research predominantly focuses on peer groups used for compensation-benchmarking purposes (Bizjak et al. 2008, 2010; Cadman and Carter 2010; Albuquerque et al. 2009; Faulkender and Yang 2010). The purpose of compensation benchmarking

is to gauge the appropriate *level* of compensation, while the purpose of RPE is mainly to employ external performance targets (i.e., peer performance) to filter out common risk. Selecting better (worse) performing firms as benchmarking (RPE) peers is expected to increase executive compensation. In additional analysis, we find that, for firms disclosing both benchmarking peers and RPE peers, peers used *solely* for RPE (benchmarking) purposes exhibit higher (lower) performance comovement with the firms of interest and have worse (better) stock performance than unselected peers. These contrasting results indicate different considerations in selecting RPE peers versus compensation-benchmarking peers. Our study thereby fills the void in the extant literature that has paid little attention to RPE peers.

We review the related literature in Section II. Section III describes the SEC's new disclosure rules, our sample-selection procedures, and descriptive statistics on RPE use for S&P 1500 firms. In Section IV, we use the implicit approach to evaluate RPE use, where we incorporate explicit RPE contract details. Section V examines firms' decisions to explicitly use RPE in executive compensation contracts, and Section VI investigates the selection of RPE peers. We conclude in Section VII.

II. RELATED LITERATURE

Literature on RPE Use

When common exogenous shocks affect the performance of multiple agents, economic theory predicts that an incentive scheme based on relative performance is superior to an incentive scheme based on individual performance if the agent is sufficiently risk-averse and common uncertainty is high (Lazear and Rosen 1981; Holmstrom 1982; Green and Stokey 1983). Since relative performance reduces the effect of common shocks on performance, RPE allows the risk-averse agent to bear less risk and the principal to better evaluate and motivate the agent's effort. Besides this risk-sharing benefit, tournament theory proposes potential competition benefits of RPE. Under certain circumstances, an incentive scheme based on relative performance can result in higher performance compared to an incentive scheme based on individual performance (Nalebuff and Stiglitz 1983; Hannan et al. 2008).

Despite its theoretical appeal, RPE involves potential costs. Heterogeneity among agents can dilute the benefits of RPE in a standard tournament (Lazear and Rosen 1981). If agents have unequal chances to win the prize given the same level of effort, a tournament can induce disadvantaged agents to shirk (O'Keefe et al. 1984) and might distort agents' risk choices (e.g., Rosen 1986; Knoeber and Thurman 1994; Hvide 2002). Prior studies also suggest that RPE contracts create adverse incentives for agents to sabotage peer performance, collude with peers, and/or choose inappropriate peer groups (Dye 1984; Gibbons and Murphy 1990; Aggarwal and Samwick 1999; Murphy 2001). Moreover, empirical evidence is consistent with competitive environments and certain executive attributes (such as self-hedging ability and outside employment opportunities) reducing the benefits of RPE in incentive contracting (e.g., Aggarwal and Samwick 1999; Garvey and Milbourn 2003; Rajgopal et al. 2006). The theory and evidence cited above provide potential explanations for a lack of consistent empirical evidence supporting the use of RPE in executive compensation contracts (see Murphy [1999] and Prendergast [1999] for reviews of related literature).

While the majority of empirical studies employ the implicit approach to test RPE use, a few studies have examined *explicit* RPE use in executive compensation contracts (Murphy 1999, 2001; Bannister and Newman 2003; Carter et al. 2009). Murphy (1999) examines 177 large U.S. firms included in the 1997 Towers Perrin survey, and Bannister and Newman (2003) examine proxy disclosures of 160 Fortune 250 firms in fiscal years 1992 and 1993. Both studies provide descriptive statistics about explicit RPE contracts and claim that a lack of empirical support for RPE use could be due to incorrect assumptions and model misspecifications in implicit tests. We investigate

this claim in Section IV. A more recent study by Carter et al. (2009) examines explicit RPE use in performance-vested equity grants for a sample of 129 U.K. firms. Their analysis of firms' choice between RPE use and non-RPE use yields little empirical support for related economic theories.³ Their U.K. evidence might not generalize to the U.S. setting, however, because performance-vested equity plans are more prevalent in the U.K., mainly due to external pressures from institutional investors (Carter et al. 2009; Ferri 2009).

Literature on RPE Peer Selection

The selection of RPE peers has significant implications for the costs and benefits of RPE, but few theoretical models specifically analyze the selection process. Nevertheless, theories on RPE use provide guidance on how to select RPE peers efficiently. To better shield executives from common exogenous shocks (Lazear and Rosen 1981; Holmstrom 1982; Green and Stokey 1983), selected peers should bear greater common risk with the RPE firms. Moreover, in a tournament, selecting agents with similar ability can reduce potential inefficiencies induced by unequal contests, such as shirking and suboptimal risk choices (e.g., Lazear and Rosen 1981; Hvide 2002). Hence, RPE firms should select peers that exhibit similar ability as themselves.

On the other hand, executive compensation practices are often the product of "managerial power," which allows rent extraction from shareholders (Bebchuk and Fried 2003). In the context of RPE, inefficient peer selection could involve deliberate selection of poorly performing firms as RPE peers to inflate peer-adjusted firm performance and thus allow executives to reap excess compensation (Gibbons and Murphy 1990; Murphy 2001). Another line of research proposes that executive compensation practices reflect symbolic considerations (Westphal and Zajac 1994; Zajac and Westphal 1995), suggesting that RPE firms prefer visible and well-established peers to facilitate the justification of their choices to external constituencies.

Due to limited disclosures about RPE peer groups, prior empirical studies generally assume that RPE peers comprise all firms in the same industry as the RPE firms. Albuquerque (2009) points out that RPE peer selection should consider both industry and size effects, and reports stronger evidence of RPE use when peers are matched on both industry and size, as opposed to industry only. Dikolli et al. (2010) examine the consequences of the implicit test's departure from the details in explicit RPE contracts. They demonstrate analytically that an empiricist's inability to observe the board's actual peer selection and aggregation of peer performance creates measurement errors that can significantly bias against detecting RPE use. Their theoretical proposition is supported by simulation results. These studies highlight the importance of peer-group composition in drawing inferences from traditional implicit tests.

III. INSTITUTIONAL BACKGROUND, SAMPLE, AND DESCRIPTIVE STATISTICS

New Disclosure Rules on Executive Compensation

Amid increased media and shareholder scrutiny of executive pay practices, the SEC amended its rules on executive compensation disclosures in 2006 to increase the transparency of executive compensation contracts. The new disclosure rules were initially proposed in January 2006 and are effective for fiscal years ending on or after December 15, 2006 (see Appendix A for a timeline related to the new disclosure rules). Prior to 2006, proxy disclosures on the details of RPE contracts in the U.S. had been voluntary (Byrd et al. 1998; Carter et al. 2009). Under the new disclosure rules, firms are required to provide a "Compensation Discussion and Analysis" (CD&A) report in their proxy statements, in which firms must provide a detailed description of the

³ Specifically, Carter et al. (2009) find no evidence that economic determinants of RPE use (such as common risk exposure) affect firms' decisions to use RPE in performance-vested equity grants as predicted by economic theories.

process used to set performance targets and an evaluation of how the performance targets translate into an objective compensation determination.⁴ This setting allows us to collect detailed data on explicit RPE contracts for a large sample of U.S. firms.

Mandatory disclosures of RPE use, however, could alter firms' decisions to use RPE and/or the implementation of RPE contracts. To the extent that the new disclosure rules attract public attention concerning the executive pay-setting process, firms could choose to restrain rent-seeking activities and thus potentially improve the compensation-contracting efficiency. Consequently, our findings might understate (overstate) the extent of rent seeking (efficient contracting) in periods prior to this new disclosure regulation.

Sample Selection

We begin with the S&P 1500 firms identified using Compustat's annual file for fiscal year 2006. We retrieve these firms' first proxy statements filed under the new executive compensation disclosure regime.⁵ Our initial sample consists of 1,419 firms with available proxy statements.

We determine the use of RPE in executive compensation contracts by reading CD&A reports. Specifically, if the firm states that at least one component of executive compensation (e.g., annual bonus, restricted stock, stock option) is determined based on firm performance relative to a group of peers, the firm is identified as a RPE firm; otherwise, the firm is identified as a non-RPE firm. We take particular care to avoid misclassifying compensation-benchmarking firms as RPE firms. Specifically, if a firm uses peer groups to solely benchmark the *level* of pay without using peer groups to determine the *performance target* when evaluating firm's performance, then the firm is coded as a non-RPE firm. Appendix B provides detailed coding criteria and representative proxy disclosures about RPE contracts in our sample. Our initial sample includes 361 RPE firms and 1,058 non-RPE firms. When firms indicate that RPE is used in setting executive pay, they usually disclose details of the RPE contracts, such as the composition of the RPE peer group and specific performance metrics used for RPE.⁶ We collect these detailed RPE contract disclosures as described in the CD&A section.

Finally, we obtain CEO compensation data from ExecuComp, financial data from Compustat, stock return data from CRSP, and analysts' target price forecasts from I/B/E/S.

⁴ SEC final rule 33-8732a, Item 402(b) (2) (v-vi) states, "In particular, firms need to disclose material information on what specific items of corporate performance are taken into account in setting compensation policies and making compensation decisions... how specific forms of compensation are structured and implemented to reflect these items of the registrant's performance, including whether discretion can be or has been exercised (either to award compensation absent attainment of the relevant performance goal(s) or to reduce or increase the size of any award or payout), identifying any particular exercise of discretion, and stating whether it applied to one or more specified named executive officers or to all compensation subject to the relevant performance goal(s)." Nontransparent or vague disclosures of RPE use may attract SEC attention to be publicized in the SEC's comment letters (see, e.g., the SEC's comment letter on Blackbaud Inc. proxy filings dated October 28, 2008).

⁵ Most firms are subject to the new disclosure rules in fiscal year 2006. Since the new rules are effective for fiscal years ending on or after December 15, 2006, and firms are required to file their proxy statements within 120 days after the fiscal year-end, some firms are not subject to the new rules until fiscal year 2007. In our initial sample of 1,419 firms, 37 firms fall into this category, among which 8 firms use RPE. For these firms, we collect information on the use of RPE from fiscal year 2007 proxy statements. For ease of exposition, we use 2006 to refer to the first year under the new disclosure rules throughout the paper.

⁶ Firms occasionally comment favorably on the principle of RPE, but do not disclose details on the use of RPE in setting compensation (e.g., no disclosure about the compensation components and performance metrics used for RPE and no details on the RPE peer group; see the last example in Appendix B). We classify these firms as non-RPE firms (66 firms), as under the SEC's new disclosure regime, firms are required to disclose the use and details of RPE contracts if they actually use RPE in setting executive compensation. Our results on firms' decisions to use RPE are qualitatively similar if we exclude these 66 non-RPE firms from the analysis (untabulated).

Descriptive Statistics

Table 1 provides descriptive evidence on the frequency of RPE use and on the main features of RPE contracts. Panel A reports that 361 (about 25 percent) of the sample firms use RPE at least to some extent in determining executive compensation. For comparison, Bannister and Newman (2003) report that 28 percent of their sample firms (which include 160 firms listed on the Fortune 250 index) use RPE over the 1992–1993 period. For S&P 500 firms, which have comparable size to Bannister and Newman's (2003) sample, we find a higher frequency of RPE use (about 38 percent), which reflects increased RPE disclosure under the SEC's new disclosure regime and/or more widespread use of RPE since the early 1990s.⁷

The fraction of RPE firms appears small compared to a theoretical prediction that RPE should be a widespread practice. We offer two explanations. First, we are only able to observe firms pre-committing to a formulaic explicit RPE contract, thereby excluding firms that use RPE implicitly through boards' subjective discretion (e.g., Ferri 2009). Second, as mentioned in Section II, RPE also involves costs, so that firms are apt to use RPE only when the perceived benefits of RPE sufficiently outweigh the costs associated with using RPE.

Table 1, Panel B reports the use of RPE by the type of compensation plans and performance metrics. As can be seen, firms most often apply RPE solely to equity-based compensation plans (214 firms, or 59 percent) and less frequently to cash compensation plans (82 firms, or 23 percent). This finding stands in contrast to the evidence reported by Murphy (1999), based on a survey of 177 large U.S. firms in 1997 that RPE is sometimes used in annual bonus plans but rarely in equity-based compensation plans. Our finding also indicates heavier use of RPE toward equity-based compensation as compared with Bannister and Newman (2003), who report that 44 percent of their sample firms use RPE for long-term plans (including restricted stock, stock options, and long-term incentive plans), based on 160 Fortune 250 firms in the early 1990s. We also find a small percentage of firms (65 firms, or 18 percent) that use RPE in setting both equity-based compensation and cash compensation.

Turning to performance metrics used in RPE contracts, Panel B of Table 1 shows that a majority of RPE firms (206 firms, or 57 percent) rely solely on price-based performance metrics (including stock returns and common shareholder wealth) when implementing RPE. Accounting performance metrics (such as return on equity and earnings per share) are also common in RPE contracts (82 firms, or 23 percent). Price-based RPE metrics are primarily used in equity-based compensation plans (150 firms out of 206 firms, or 73 percent), while cash compensation plans more frequently use accounting-based RPE metrics (35 firms out of 82 firms, or 43 percent). Among the remaining RPE firms, 60 firms (17 percent) use both price-based and accounting performance metrics for RPE purposes. An even smaller percentage of firms (13 firms, or 4 percent) use nonfinancial performance metrics (such as customer satisfaction and market share) in RPE contracts. The infrequent use of nonfinancial metrics potentially reflects difficulties in obtaining peer firms' nonfinancial information (e.g., Ferri 2009). We list commonly used RPE performance metrics in Table 1, Panel C.

Firms that use RPE in setting equity-based compensation can employ RPE in a number of ways. Table 1, Panel D shows that a significant portion of RPE firms (about 43 percent) use RPE

⁷ To better understand the effect of the new disclosure rules on firms' disclosures of RPE use, we compare S&P 500 firms' RPE disclosures before and after the new disclosure rules. For the 490 S&P 500 firms in our sample, the percentage of RPE firms in 2005 is about 31 percent (153 firms), compared to 38 percent (185 firms) in 2006. This result appears consistent with more firms adopting RPE in anticipation of more public scrutiny under the new disclosure rules and is also consistent with increased disclosure of RPE use under the new disclosure regime.

TABLE 1
Descriptive Statistics of Relative Performance Evaluation Use in Executive Compensation Plans

Panel A: Relative Performance Evaluation Use^a

	Number of Firms	Percentage (%)
RPE firms	361	25.44
Non-RPE firms	1,058	74.56
Total	1,419	100.00

Panel B: Relative Performance Evaluation Use by Type of Compensation Plan and Performance Metrics^b

	Price-Based Metrics	Accounting Metrics	Price-Based and Accounting Metrics	Others	Total
Equity-based compensation plan only					
Restricted stock	126	26	17	3	172
Stock option	3	1	1	0	5
Restricted stock and stock option awards	21	7	6	3	37
Subtotal	150	34	24	6	214
Cash compensation plan only	32	35	12	3	82
Equity-based and cash compensation plans	24	13	24	4	65
Total	206	82	60	13	361

Panel C: Commonly Used Price-Based and Accounting Performance Metrics for Relative Performance Evaluation^c

	Number of Firms	Percentage (%)
Stock returns	266	73.68
Return on equity	50	13.85
Growth of earnings	42	11.63
Earnings	39	10.81
Growth of sales	24	6.65
Return on asset	14	3.88
Cash flow	13	3.60

Panel D: Details of Relative Performance Evaluation Use in Equity-Based Compensation Plan^d

	Number of Firms	Percentage (%)
RPE determines		
Vesting condition of performance share units (type 1)	120	43.01
Size of restricted stocks and stock options (type 2)	94	33.69
Vesting condition of restricted stocks and stock options (type 3)	61	21.86
Both type 1 and type 3	2	0.72
Both type 2 and type 3	2	0.72
Total	279	100.00

(continued on next page)

Panel E: Relative Performance Evaluation Use by Type of Compensation Plan and Peer Group Choice^a

	Self-Selected Peers	Market/Industry Index	Market/Industry Index and Self-Selected Peers	Total
Equity-based compensation plan only				
Restricted stock	91	72	9	172
Stock option	4	1	0	5
Restricted stock and stock option awards	21	10	6	37
Subtotal	116	83	15	214
Cash compensation plan only	51	25	6	82
Equity-based and cash compensation plans	42	19	4	65
Total	209	127	25	361

Panel F: Number of Peers in the Self-Selected Peer Groups and Industry Similarity and S&P 1500 Identity for Self-Selected Peers^a

	Mean	Standard Deviation	25%	Median	75%	n
Number of self-selected peers	14.694	8.721	9	13	18	232
Percentage of self-selected peers in the same two-digit SIC industry as the RPE firms (%)	60.92	33.39	34.85	66.67	91.67	232
Percentage of self-selected peers in the same three-digit SIC industry as the RPE firms (%)	47.92	34.93	14.84	45.45	83.33	232
Percentage of self-selected peers in the S&P 1500 index (%)	70.37	22.42	58.58	74.60	86.36	232
Percentage of self-selected peers in the same S&P 1500 sub-index as the RPE firms (%)	45.11	27.77	23.67	40.00	67.71	232

^a RPE (Non-RPE) firms include 361 (1,058) S&P 1500 firms listed in the Compustat annual file for 2006 that explicitly disclose (do not mention) the use of relative performance evaluation in executive compensation plans in their first annual proxies filed under the SEC's 2006 executive compensation disclosure rules. We are unable to find proxy statements for 81 S&P 1500 firms.

^b "Equity-based compensation plan only" category includes firms that use relative performance evaluation in setting equity-based compensation (such as restricted stock and stock option). "Cash compensation plan only" category includes firms that use relative performance evaluation in setting cash compensation (such as annual bonus and long-term incentive payout). "Equity-based and cash compensation plans" category includes firms that use RPE in both equity-based and cash compensation plans.

"Price-Based Metrics" category includes firms that use price-based performance metrics (such as stock returns and shareholder wealth) to implement relative performance evaluation in compensation plans. "Accounting Metrics" category includes firms that use accounting performance metrics (such as return-on-equity, earnings per share, earnings growth) to implement relative performance evaluation in compensation plans. "Price-Based and Accounting Metrics" category includes firms that use both price-based and accounting performance metrics to implement relative performance evaluation in compensation plans. "Others" category includes firms that use nonfinancial performance metrics (such as customer satisfaction and market share) to implement relative performance evaluation in compensation plans. See Panel C for a more detailed breakdown of performance metrics.

^c The percentages sum to more than 100 percent as some firms employ multiple performance metrics when applying relative performance evaluation in compensation plans.

^d "Vesting condition of performance share units" category includes firms that use RPE to determine the vesting of "performance share units (PSU)." Performance shares are similar to restricted stocks in that they fully vest at the end of a certain performance period starting with the grant year (usually three years). However, the range of vesting amount for PSU is usually between 0 percent and 200 percent (sometimes 150 percent), which is different from that for restricted stock units (RSU) capped at 100 percent. "Size of restricted stocks and stock options" category includes firms that use RPE to determine the size of granted restricted stocks and stock options. "Vesting condition of restricted stocks and

(continued on next page)

stock options" category includes firms whose stock options or RSUs are vested conditioned on achievement of RPE goals. This category also includes firms for which the vesting schedule of stock options or restricted stock is accelerated if the company achieves a certain RPE target.

- "Self-Selected Peers" category includes firms that chose individual firms as the peer groups to implement relative performance evaluation in compensation plans. "Market/Industry Index" category includes firms that chose published market/industry indices as the peer groups to implement relative performance evaluation in compensation plans. "Market/Industry Index and Self-Selected Peers" includes firms that chose both published market/industry indices and individual firms as the peer groups to implement relative performance evaluation in compensation.
- ^f Two firms mentioned the use of self-selected peers in implementing relative performance evaluation in executive compensation plans, but did not disclose the composition of the RPE peer group, which reduces the sample size to 232 firms. On a firm level, there are 39 (25) RPE firms whose self-selected peer firms all come from the same two-digit (three-digit) SIC industry; there are 25 (10) RPE firms whose self-selected peer firms all come from the S&P 1500 index (the same S&P 1500 sub-index) as the RPE firms.

to determine the vesting of performance share units.⁸ A number of firms (about 34 percent) use RPE to determine the size of restricted stock and stock option grants. Other firms (about 22 percent) link the achievement of RPE performance target to vesting conditions of restricted stocks and stock options. These statistics indicate that the use of RPE in equity grants now commonly involves vesting conditions among U.S. firms.

Finally, we examine the choice and composition of RPE peer groups. Table 1, Panel E reports that a majority of RPE firms (209 firms, or 58 percent) use self-selected RPE peer groups, and about one-third of RPE firms (127 firms, or 35 percent) employ published market or industry indices. A small number of RPE firms (25 firms, or 7 percent) use both published indices and self-selected peer groups in RPE contracts. The tendency to use self-selected peer groups (as opposed to published indices) does not seem to differ notably across equity-based compensation plans and cash compensation plans.

Table 1, Panel F shows that the mean (median) number of peers in self-selected peer groups is about 15 (13) firms, in contrast to a mean (median) of 60 (65) firms in these firms' corresponding two-digit SIC industry. Moreover, only 61 percent (48 percent) of self-selected peers operate in the same two-digit (three-digit) SIC industry as the RPE firms. At the firm level, there are only 39 (25) RPE firms whose self-selected peers all come from the same two-digit (three-digit) SIC industry. These statistics reinforce the concern that the key assumption underlying implicit tests—that the RPE peer group consists of all firms from the same industry—is inaccurate.⁹ Finally, about 70 percent of self-selected peers belong to the S&P 1500 index, and about 45 percent share the same S&P sub-index (S&P 500, Mid-Cap 400, and Small-Cap 600) as the RPE firms. At the firm level, there are 25 (10) RPE firms whose self-selected peers all come from the S&P 1500 index (the same S&P 1500 sub-index).

Table 2 provides descriptive statistics on firm characteristics, management attributes, and corporate governance for the sample firms. As shown, firms that disclose RPE use in executive compensation contracts differ from non-RPE firms along many dimensions, including competitive

⁸ Performance shares are similar to restricted stocks in that they fully vest at the end of a specified performance period, starting with the grant year (usually three years). However, the vesting amount for performance shares ranges between 0 percent and 200 percent (see the example of PG&E Corp in Appendix B), whereas that for restricted stocks is capped at 100 percent.

⁹ A possible reason for a RPE firm to select peers outside its (primary) industry is that the RPE firm follows diversified business strategies and operates with multiple segments, which requires a diversified peer group to appropriately benchmark firm performance. We find that, of the 232 RPE firms that have self-selected peer groups, 157 firms (about 68 percent) operate with multiple business segments. Multiple-segment RPE firms have roughly 61 percent (45 percent) of peers selected from the same two-digit (three-digit) SIC industry, while single-segment RPE firms have about 72 percent (65 percent) of peers selected from the same two-digit (three-digit) SIC industry.

TABLE 2
Descriptive Statistics of the Sample

	Non-RPE Firms (n = 1,058)		RPE Firms (n = 361)			
	Mean	Median	Mean	Median	Mean Differences	Median Differences
Firm Characteristics						
<i>Common_Risk</i>	0.285	0.257	0.384	0.370	−0.099***	−0.113***
<i>lSize_Rkadjl</i>	5.348	5.081	6.325	5.969	−0.977***	−0.888***
<i>lDiversity_Rkadjl</i>	0.019	0.007	0.024	0.015	−0.005***	−0.008***
<i>lReturn_Rkadjl</i>	0.049	0.024	0.040	0.022	0.009**	0.002
<i>Industry_Concentration</i>	0.124	0.085	0.110	0.072	0.014**	0.013***
<i>BM</i>	0.600	0.602	0.682	0.692	−0.082***	−0.090***
<i>Size</i>	7.654	7.466	8.484	8.355	−0.830***	−0.889***
<i>ROA_Indadj</i>	0.014	0.000	0.005	0.000	0.009	0.000
<i>Return_Indadj</i>	0.017	−0.007	0.019	−0.001	−0.002	−0.006
CEO Attributes						
<i>CEO_Wealth</i>	8.972	9.868	9.435	10.011	−0.463	−0.143
<i>CEO_Age</i>	55.882	56.000	56.053	56.000	−0.171	0.000
Corporate Governance						
<i>Top5_Instown</i>	0.392	0.372	0.374	0.353	0.018**	0.019**
<i>Activist_Instown</i>	0.028	0.028	0.028	0.029	0.000	−0.001**
<i>CEO/Chair</i>	0.601	1.000	0.659	1.000	−0.058*	0.000
<i>Board_Independence</i>	0.817	0.857	0.857	0.889	−0.040***	−0.032***
<i>Board_Size</i>	8.957	9.000	10.518	10.000	−1.561***	−1.000***
<i>CompConsultant</i>	0.798	1.000	0.904	1.000	−0.106***	0.000

*, **, *** Indicate significance at less than the 10 percent, 5 percent, and 1 percent levels, respectively, based on two-tailed t-tests (z-tests) on mean (median) differences.

The sample includes S&P 1500 firms listed in the Compustat annual file for 2006 that have available proxy statements from the SEC's EDGAR database. RPE (Non-RPE) firms include 361 (1,058) S&P 1500 firms that explicitly disclose (do not mention) the use of relative performance evaluation in executive compensation plans in their first annual proxies filed under the SEC's 2006 executive compensation disclosure rules.

Variables are measured over or at the end of fiscal year 2005 (2006) if a firm's first proxy filed under the SEC's 2006 executive compensation disclosure rules is for fiscal year 2006 (2007). Firm characteristics and CEO attributes (except *Industry_Concentration* and *CEO_Age*) are winsorized at the top and bottom one percentiles. Number of observations varies depending on data availability.

Variable Definitions:

Common_Risk = proportion of firm-level stock return variance that is explained by value-weighted industry stock returns, measured by the R^2 from regressing firm-level stock returns on value-weighted industry stock returns over the prior 36 months;

lSize_Rkadjl = natural logarithm of the absolute difference between the firm's market value of equity and the median market value of equity for the firm's corresponding decile;

lDiversity_Rkadjl = natural logarithm of the absolute difference between the firm's segment concentration and the median segment concentration for the firm's corresponding decile. Segment concentration is the sum of the squares of each segment's sales as a percentage of the total firm sales;

lReturn_Rkadjl = absolute difference between the firm's annual stock returns and the median annual stock returns for the firm's corresponding decile;

Industry_Concentration = sum of the squares of the market shares of the firms' sales within each two-digit SIC industry;

BM = book value of assets divided by the sum of the market value of equity and book value of liabilities;

Size = natural logarithm of the market value of equity;

(continued on next page)

TABLE 2 (continued)

<i>ROA_Indadj</i>	= return-on-assets minus the median return-on-assets for the same industry (two-digit SIC code);
<i>Return_Indadj</i>	= buy-and-hold annual stock returns minus the median buy-and-hold annual stock returns for the same industry (two-digit SIC code);
<i>CEO_Wealth</i>	= natural logarithm of the value of equity (including both stocks and stock options) held by the CEO;
<i>CEO_Age</i>	= age of the CEO;
<i>CompConsultant</i>	= 1 if the firm uses compensation consultant in the pay-setting process (as disclosed in their first annual proxies filed under the SEC's 2006 executive compensation disclosure rules), and 0 otherwise;
<i>Top5_Instown</i>	= stock ownership by the top five institutions as a percentage of total institutional ownership;
<i>Activist_Instown</i>	= percentage of holdings by activist institutions as defined by Cremers and Nair (2005);
<i>CEO/Chair</i>	= 1 if the CEO serves as Chairman of the Board, and 0 otherwise;
<i>Board_Independence</i>	= percentage of independent directors serving on the board; and
<i>Board_Size</i>	= number of directors serving on the board.

environment, growth opportunities, common risk exposure, and market capitalization. In addition, RPE firms have different ownership structure and board attributes than non-RPE firms, and more frequently employ compensation consultants when setting executive compensation. We consider these differences when conducting multivariate analyses in Section V.

IV. INCORPORATING EXPLICIT RPE CONTRACT DETAILS IN IMPLICIT TESTS

Prior empirical studies generally employ an implicit approach to test for the prevalence of RPE use *without* referring to explicit RPE contracts. In this section, we examine the extent to which incorporating details of explicit RPE contracts affects inferences drawn under the implicit approach. Our approach, while incorporating important features of RPE contracts, still relies on assumptions regarding how RPE is actually utilized in setting executive pay.¹⁰ Nevertheless, such evidence would highlight limitations of implicit tests and underscore the importance of considering explicit RPE contract details in drawing conclusions about RPE use.

We start by following Albuquerque (2009)'s implicit approach with one exception. Albuquerque (2009) included CEO fixed effects for her multiple-year sample period of 1992–2005. We estimate the following equation using S&P 1500 firms, where we limit the analysis to 2006 because we only collect information about explicit RPE contracts for 2006:

$$CEOPay_t = \alpha_0 + \alpha_1 FirmPerf_t + \alpha_2 PeerPerf_t + \alpha_3 ControlVariables_t + \varepsilon_t. \quad (1)$$

In Equation (1), the dependent variable, *CEOPay*, is CEO total compensation, measured as the sum of salary, bonus, the fair value of stock option awards and restricted stock awards (per SFAS 123R), the change in deferred compensation, non-equity incentive plan compensation, and other compensation, in thousands of dollars (i.e., TDC1 from ExecuComp). To mitigate skewness in CEO total compensation, we follow prior studies and use its natural logarithm in the estimation (e.g., Murphy 1999; Albuquerque 2009). The independent variables, *FirmPerf_t* and *PeerPerf_t*, are 12-month buy-and-hold stock returns during the fiscal year for the firm and its RPE peer group, respectively. Albuquerque (2009) proposes firm size as an important factor in RPE peer selection, while prior implicit tests utilize misspecified peer groups by combining same-industry firms with-

¹⁰ For instance, we assume that all firms use stock return as the RPE-based performance metric and the median performance among peers determines the hurdle rate in achieving RPE-based performance targets.

out considering firm size. We follow Albuquerque's (2009) recommendation and measure *PeerPerf_i*, based on the median stock return for firms in the same size quartile and same industry. If firms filter out peer performance in setting CEO pay, then we expect to find $\alpha_2 < 0$.

Equation (1) also controls for firm characteristics and governance attributes that can influence CEO pay, including firm size, growth options, CEO tenure, regulated industry membership, idiosyncratic variance, CEO/Chair duality, CEO stock ownership, and whether the CEO is involved in an interlocking relationship. To account for potential outliers, we winsorize non-dichotomous variables at top and bottom one percentiles (except that idiosyncratic variance is winsorized at the top one percentile only). We further include industry indicator variables to account for variations in CEO pay at the industry level.

Table 3 presents the results from estimating Equation (1). The first column reports results for our full sample of S&P 1500 firms. The sample size is smaller than that reported in Table 1, Panel A (reduced by 291 firms) due to (1) additional sample restrictions (excluding new CEOs who take the position in the middle of 2006 and excluding firms reporting negative total compensation), and (2) missing values on control variables. As shown, the coefficient on industry-size-matched peer returns is insignificant, inconsistent with RPE use.¹¹ The lack of evidence in support of RPE use could result from infrequent use of RPE in 2006, as only 25 percent of our sample firms explicitly disclose the use of RPE in setting executive pay. If the infrequent use of RPE underlies the insignificant result above, then we expect to observe stronger evidence on RPE use for RPE firms than for non-RPE firms. Columns (2) and (3) report the results for non-RPE firms and RPE firms, respectively. We find that for both non-RPE firms and RPE firms, the coefficient on industry-size matched peer returns is insignificant. Hence, even among firms explicitly disclosing RPE use, the implicit test provides no evidence that these firms filter out peer performance when setting CEO pay. These results indicate that the infrequent use of RPE is unlikely to be the dominant reason for the insignificant evidence on RPE use in the implicit test.

To further refine the implicit test, we consider disclosed RPE contract details. First, we incorporate RPE peer-group composition by measuring *PeerPerf_i*, based on the median stock return of self-selected RPE peers as disclosed in firms' 2006 proxy statements. Table 3, Column (4) shows that the coefficient on self-selected peer returns is significantly negative, consistent with firms using these self-selected peers for RPE purposes when setting CEO pay. The contrasting results across Columns (3) and (4) highlight the importance of peer-group composition in testing for RPE use.¹² The implicit test is likely to produce misleading results due to inaccurate identification of RPE peers employed in the pay-setting process.

Second, we consider performance-based vesting, a design feature of RPE contracts that has gained popularity in the United States in recent years (Gerakos et al. 2007). The traditional implicit tests use *current* performance for RPE firms and their peers to explain executive pay. However, as Table 1, Panel D documents, firms often link RPE-based performance targets to vesting conditions of equity grants, so that *future* performance determines the value of equity grants (Gerakos et al. 2007; Carter et al. 2009). Under SFAS 123R, firms are required to report the fair value of performance-vested equity grants based on estimates of future performance for both

¹¹ This result appears to be inconsistent with Albuquerque's (2009) main finding of a negative coefficient on industry-size-matched peer returns. This inconsistency is driven by our different sample period: Albuquerque's (2009) sample period ranges from 1992 to 2005, while our sample period is limited to 2006.

¹² The sample size under Column (3) is larger than the sample size under Column (4) because Column (4) focuses attention on RPE firms disclosing self-selected RPE peers. Results are qualitatively similar if we restrict to RPE firms disclosing self-selected RPE peers in Column (3). The sample size under Column (4) is smaller than the sample size reported in Table 1, Panel E (i.e., 209 firms with self-selected RPE peer groups) because of additional sample restrictions (excluding new CEOs who take the position in the middle of 2006 and excluding firms reporting negative total compensation) and missing values on control variables.

TABLE 3
Implicit Test of RPE Use With and Without Incorporating Explicit RPE Contract Details

	Predicted Sign	Without Incorporating Explicit RPE Contract Details		Incorporating Explicit RPE Contract Details	
		RPE Firms with Self-Selected Peers		Linked to Current Peer Performance	
		Full Sample (S&P 1500)	Non-RPE Firms	RPE Firms	Linked to Future Peer Performance
		(1)	(2)	(3)	(6)
		Coefficient (p-value)	Coefficient (p-value)	Coefficient (p-value)	Coefficient (p-value)
Independent Variables					
<i>Firm Stock Return</i>	+	0.329 (0.001)	0.273 (0.013)	0.928 (<0.001)	0.759 (0.077)
<i>Peer Stock Return (Industry/Size)</i>	-	0.284 (0.413)	0.053 (0.898)	0.604 (0.346)	0.694 (0.039)
<i>Peer Stock Return (Self-Selected Peers)</i>	-				
Control Variables					
<i>Firm Size (Sales)</i>	+	0.468 (<0.001)	0.478 (<0.001)	0.427 (<0.001)	0.450 (<0.001)
<i>Growth Option</i>	+	0.080 (0.005)	0.065 (0.033)	0.225 (<0.001)	0.241 (0.006)
<i>CEO Tenure</i>	+	-0.074 (0.040)	-0.056 (0.182)	-0.136 (0.056)	-0.074 (0.642)
<i>Regulation Industry</i>	-	-0.245 (0.097)	-0.204 (0.421)	-0.472 (0.379)	-1.042 (0.018)
<i>Idiosyncratic Variance</i>	+	0.871 (0.411)	0.427 (0.720)	3.890 (0.105)	2.509 (0.566)

(continued on next page)

TABLE 3 (continued)

	Without Incorporating Explicit RPE Contract Details		Incorporating Explicit RPE Contract Details			
	RPE Firms with Self-Selected Peers		Linked to Current Peer Performance		Linked to Future Peer Performance	
	All		(5)		(6)	
	(1)	(2)	(3)	(4)	(5)	(6)
	Full Sample (S&P 1500)	Non-RPE Firms	RPE Firms	Coefficient (p-value)	Coefficient (p-value)	Coefficient (p-value)
Predicted Sign	Coefficient (p-value)	Coefficient (p-value)	Coefficient (p-value)	Coefficient (p-value)	Coefficient (p-value)	Coefficient (p-value)
<i>CEO/Chair</i>	0.233 (<0.001)	0.201 (0.001)	0.375 (<0.001)	0.468 (<0.001)	0.557 (0.006)	0.578 (0.003)
<i>CEO Ownership</i>	-0.068 (0.292)	-0.073 (0.315)	0.061 (0.709)	0.154 (0.461)	0.104 (0.688)	0.095 (0.813)
<i>Interlock</i>	-0.938 (0.026)	-1.436 (<0.001)	0.585 (0.633)	0.403 (0.707)	0.413 (0.713)	NA
Intercept	4.446 (<0.001)	4.714 (<0.001)	4.231 (<0.001)	4.477 (<0.001)	4.971 (<0.001)	5.009 (<0.001)
Industry fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R ²	0.550	0.536	0.659	0.728	0.741	0.805
Number of observations	1,128	837	291	184	95	89

The full sample includes 1,128 S&P 1500 firms from the ExecuComp database for the year 2006 merged with Compustat and CRSP databases. The standard errors are heteroscedasticity-consistent using the Huber-White correction and are clustered by firm. The reported p-values in parentheses are two-tailed. The dependent variable, *CEOPay*, is the natural logarithm of total CEO compensation for 2006 in thousands of dollars (TDC1).

Variable Definitions:

Firm Stock Return = 12-month buy-and-hold stock returns for 2006;

Peer Stock Return = measured in two ways: *Peer Stock Return (Industry/Size)* is the median 12-month buy-and-hold stock returns for 2006 (excluding own-firm performance)

(continued on next page)

TABLE 3 (continued)

for the industry and size-matched peer group (see Albuquerque (2009) for details); *Peer Stock Return (Self-Selected Peers)* is the median 12-month buy-and-hold stock returns for 2006 for the self-selected peer group by RPE firms;

Firm Size (Sales) = natural logarithm of sales for 2005;

Growth Option = ratio of the market value of equity to the book value of assets at the beginning of 2006;

CEO Tenure = natural logarithm of the number of years since the CEO assumed office;

Regulation Industry = an indicator variable equal to 1 for firms in the gas and electric industries with SIC codes from 4900 to 4939, and 0 otherwise;

Idiosyncratic

Variance = difference between firm-level stock return variance and the industry's average return variance, calculated over the previous 36 months;

CEO/Chair = indicator variable equal to 1 if the CEO is the board chair, and 0 otherwise;

CEO Ownership = indicator variable equal to 1 if the proportion of shares held by the CEO is lower than the sample median, and 0 otherwise; and

Interlock = indicator variable equal to 1 if the CEO is involved in an interlock relationship requiring disclosure in the proxy statement, and 0 otherwise.

the RPE firms and their peers, as well as the associated likelihood of future vesting realizations. Because compensation committees' projections of future performance can differ from current performance, the use of current performance in testing for RPE use could produce low-power tests when RPE-based performance targets are linked to future performance rather than current performance. The last two columns of Table 3 examine this issue. As shown in Column (5), when RPE use is linked to current peer performance, the coefficient on self-selected peer returns is significantly negative, supporting the use of RPE. In contrast, Column (6) shows that, when RPE use is linked to future peer performance (through equity grants' vesting conditions), there is no evidence of RPE use. Hence, failure to consider this feature of RPE contracts is likely to confound empirical results based on implicit tests of RPE use.¹³

V. USE OF RPE IN EXECUTIVE COMPENSATION CONTRACTS

Prior studies have proposed a variety of factors that affect firms' decisions to use RPE in executive compensation contracts. However, these studies rely on the implicit approach and mostly examine one contextual factor at a time without controlling for alternative factors. In this section, we employ explicit disclosures on RPE use to simultaneously examine multiple factors that influence the decision to incorporate RPE into executive compensation contracts.

Determinants of RPE Use

Common Risk

Theory predicts that the risk-sharing benefit of RPE increases in the extent to which common risk affects both firm and peer performance (e.g., Holmstrom 1982).¹⁴ We thus expect firms to be more likely to use RPE when their performance is more heavily influenced by common risk. We measure common risk based on the proportion of the variance of firm-level stock returns that can be explained by industry returns, defined as the R^2 from regressing firm-level stock returns on value-weighted industry stock returns over the prior 36 months (*Common_Risk*).

Availability of Similar Peers

Tournament theory implies that including firms with unequal ability in the RPE peer group could induce executives to shirk or to pursue inefficient investments. Firms are thus less likely to use RPE if potential peers all exhibit dissimilar ability. We argue that firm size and industry diversification reflect organizational complexity, and firm performance reflects existing economic conditions, all of which affect executives' ability to compete with potential peers. To measure the availability of peers with similar ability, we first divide firms within the same two-digit SIC industry into deciles based on market value of equity, segment concentration defined as the Herfindahl-Hirschman Index of segment sales (Rose and Shepard 1997; Bushman et al. 2004), or annual stock returns. The availability of similar peers is then measured by the absolute differences in market value of equity, segment concentration, and annual stock returns, respectively, between the firm and the median of the decile to which the firm belongs (*|Size_Rkadj|*, *|Diversity_Rkadj|*, and *|Return_Rkadj|*, respectively).

¹³ Other features of RPE contracts relate to compensation components, performance metrics, hurdle rates in achieving RPE-based performance targets (e.g., median, 25 percent, or 75 percent of peer group performance), etc. Combining all of the features of RPE contracts in studying RPE use would substantially reduce our sample size and, hence, is not feasible to implement.

¹⁴ While it is desirable to filter out common shocks for better risk-sharing, some firms may deliberately choose not to use RPE to motivate executives to manage risks resulting from such exogenous shocks (e.g., hedging in oil and gas industry).

Industry Competition

Aggarwal and Samwick (1999) and Joh (1999) argue that firms facing a more competitive environment are less likely to use RPE due to the concern that RPE may encourage destructive competition. Following these studies, we use industry concentration, measured as the Herfindahl-Hirschman Index of sales within each two-digit SIC industry, as an inverse proxy for industry competition (*Industry_Concentration*).

Growth Opportunities

Murphy (2001) predicts that firms with high-growth opportunities are more likely to adopt external standards (such as peer performance) in setting executive pay because internal standards (such as historical earnings growth) provide managers of high-growth firms stronger incentives to smooth performance in response to fears of budget ratcheting (Leone and Rock 2002). We use the book-to-market ratio (*BM*) as an inverse proxy for growth opportunities.

Size

Himmelberg and Hubbard (2000) use firm size to capture CEO talent. They propose that the reservation wage for talented CEOs covaries with the state of the economy, such that firms are less likely to filter out industry and market-wide performance when compensating talented CEOs (also see Rajgopal et al. 2006). Alternatively, firm size could serve as a crude proxy for public scrutiny and shareholder concerns about executive pay practices (Bannister and Newman 2003), which may induce firms to commit to explicit RPE use. We use the market value of equity to measure firm size (*Size*).

Firm Performance

Firm performance could also proxy for CEO talent and, hence, could be negatively related with RPE use (Himmelberg and Hubbard 2000). Alternatively, firms having superior performance could prefer RPE as a convenient tool to justify higher executive pay. We include industry-adjusted operating performance (*ROA_Indadj*) and stock performance (*Return_Indadj*) to capture performance effects on RPE use.

Executive Attributes

Garvey and Milbourn (2003) argue that the theoretical benefit of RPE is diminished to the extent that agents can hedge market risk. Hence, firms are less likely to use RPE when their CEOs display greater self-hedging ability. Following Garvey and Milbourn (2003), we use the value of CEO equity holdings (*CEO_Wealth*) and CEO age (*CEO_Age*) to proxy for CEOs' self-hedging ability.

Ownership Structure and Board Attributes

Bertrand and Mullainathan (2001) document that CEO pay is positively associated with exogenous shocks and that such "pay-for-luck" is less pronounced in better-governed firms. Garvey and Milbourn (2006) further show that executive pay is more sensitive to good luck than to bad luck and that this asymmetry is more pronounced for firms with weaker governance. Since RPE can limit the pay-for-luck practice and also has risk-sharing benefit, we expect that RPE use is positively associated with the quality of corporate governance. We measure governance quality using institutional ownership concentration (*Top5_Instown*), activist institutional ownership (*Activist_Instown*) as defined in Cremers and Nair (2005), CEO/Chair duality (*CEO/Chair*), the proportion of outside directors on the board (*Board_Independence*), and the number of directors sitting on the board (*Board_Size*).

Compensation Consultant

Compensation consultants have access to proprietary information about industry-wide compensation practices and potential competitors (Cadman et al. 2010; Murphy and Sandino 2010). Since the design and implementation of RPE contracts demand specialized knowledge and expertise, the availability of professional advice from compensation consultants is likely to influence firms' decisions to use RPE. We construct an indicator variable to identify firms that use compensation consultants in setting executive pay (*CompConsultant*), based on firms' first proxy disclosures under the SEC's new disclosure rules.

Results on RPE Use

To examine firms' decisions to use RPE in executive compensation contracts, we estimate the following multivariate logistic regression:

$$\begin{aligned} \text{Prob}(RPE_t = 1) = \Phi(\alpha_0 + \alpha_1 \text{Common\_Risk}_{t-1} + \alpha_2 |\text{Size\_Rkadj}|_{t-1} + \alpha_3 |\text{Diversity\_Rkadj}|_{t-1} \\ + \alpha_4 |\text{Return\_Rkadj}|_{t-1} + \alpha_5 \text{Industry\_Concentration}_{t-1} + \alpha_6 \text{BM}_{t-1} \\ + \alpha_7 \text{Size}_{t-1} + \alpha_8 \text{ROA\_Indadj}_{t-1} + \alpha_9 \text{Return\_Indadj}_{t-1} + \alpha_{10} \text{CEO\_Wealth}_{t-1} \\ + \alpha_{11} \text{CEO\_Age}_{t-1} + \alpha_{12} \text{Top5\_Instown}_{t-1} + \alpha_{13} \text{Activist\_Instown}_{t-1} \\ + \alpha_{14} \text{CEO/Chair}_{t-1} + \alpha_{15} \text{Board\_Independence}_{t-1} + \alpha_{16} \text{Board\_Size}_{t-1} \\ + \alpha_{17} \text{CompConsultant}_t + \varepsilon_t). \end{aligned} \quad (2)$$

The dependent variable, *RPE*, is an indicator variable that equals 1 for RPE firms and 0 for non-RPE firms. The independent variables include firm characteristics, CEO attributes, and corporate governance factors that potentially influence the use of RPE.

Table 4 reports the results from estimating Equation (2). We find that firms exposed to greater common risk are more likely to use RPE, in line with the notion that RPE is more effective at removing exogenous shocks in performance evaluation when the firm shares greater common risk with its peers (Janakiraman et al. 1992). In contrast to the prediction that more intense product market competition (lower industry concentration) discourages the use of RPE (Aggarwal and Samwick 1999), we find that firms from less concentrated industries are more likely to implement RPE in compensation plans. This result, however, is consistent with the evidence in DeFond and Park (1999) that a more competitive environment is characterized by a higher degree of common risk and, hence, brings greater benefits to those who use RPE.¹⁵ We also find that firms exhibiting fewer growth opportunities (higher book-to-market) are more likely to use RPE, which contradicts Murphy's (2001) prediction but is in line with the argument in Albuquerque (2007) that RPE is less beneficial for firms with high-growth options, since peer performance is a less informative signal of external shocks for these firms.¹⁶

Furthermore, we find that larger firms are more likely to use RPE, consistent with firm size serving as a proxy for public scrutiny and shareholder concerns about executive pay. We also find that firms are less likely to use RPE when the value of CEO equity holdings is large, consistent

¹⁵ DeFond and Park (1999) examine RPE use in management turnover decisions and find stronger evidence of RPE use in highly competitive industries than in less competitive industries. Our finding that *Industry_Concentration* (inverse proxy for industry competition) is negatively related to RPE use, even after controlling for *Common_Risk*, suggests that industry concentration may capture common risk above and beyond *Common_Risk*.

¹⁶ Albuquerque (2007) argues that growth options affect a firm's common risk exposure and, hence, the informativeness of peer performance about the firm's common risk. Given we have controlled for *Common_Risk*, the positive relation between book-to-market and RPE use suggests that book-to-market may capture common risk above and beyond *Common_Risk*.

TABLE 4
Logistic Regression of Relative Performance Evaluation Use in Executive Compensation Contracts

Independent Variables	Predicted Sign	Coefficient (p-value)	Change in Probability Q1 versus Q3 Values
Firm Characteristics			
<i>Common_Risk</i>	+	1.206 (0.016)	0.088
<i>lSize_Rkadj</i>	—	−0.087 (0.231)	−0.049
<i>lDiversity_Rkadj</i>	+/−	4.712 (0.222)	0.039
<i>lReturn_Rkadj</i>	—	1.257 (0.303)	0.010
<i>Industry_Concentration</i>	+	−6.022 (0.001)	−0.097
<i>BM</i>	—	1.499 (0.005)	0.118
<i>Size</i>	+/−	0.451 (0.001)	0.026
<i>ROA_Indadj</i>	+/−	−0.423 (0.641)	−0.009
<i>Return_Indadj</i>	+/−	−0.016 (0.874)	−0.001
CEO Attributes			
<i>CEO_Wealth</i>	—	−0.168 (0.002)	−0.051
<i>CEO_Age</i>	—	0.006 (0.664)	0.014
Corporate Governance			
<i>Top5_Instown</i>	+	−0.321 (0.738)	−0.010
<i>Activist_Instown</i>	+	4.837 (0.688)	0.011
<i>CEO/Chair</i>	—	0.193 (0.322)	0.048
<i>Board_Independence</i>	+	4.614 (0.000)	0.011
<i>Board_Size</i>	+/−	0.166 (0.001)	0.071
<i>CompConsultant</i>	+	0.328 (0.052)	0.081
Intercept		−9.861 (<0.001)	
Number of RPE/Non-RPE firms		255/640	
Percent concordant/discordant		76.9/22.9	
Pseudo R^2		0.255	
Wald χ^2		118.071	
(p-value)		(<0.001)	

(continued on next page)

TABLE 4 (continued)

The sample includes 895 S&P 1500 firms listed in the Compustat annual file for 2006 that have available information on relative performance evaluation use in executive compensation contracts and available information on regression variables. The dependent variable is *RPE*, which is equal to 1 if the firm explicitly reported RPE use in executive compensation plans in its first annual proxy filed under the SEC's 2006 executive compensation disclosure rules, and 0 otherwise. See Table 2 for other variable definitions. Firm characteristics and CEO attributes (except *Industry_Concentration* and *CEO_Age*) are winsorized at the top and bottom one percentiles. The "Change in Probability" column shows the change in the probability of using RPE as the result of moving from the first to the third quartile value of the variable of interest (for indicator variables, moving from 0 to 1), holding all other variables constant at their mean values (for indicator variables, the benchmark probability is determined with the 0 value). The reported p-values in parentheses are two-tailed.

with Garvey and Milbourn (2003). Finally, firms with more independent boards are more likely to use RPE, consistent with stronger internal governance favoring the use of RPE, and firms that hire compensation consultants are also more likely to use RPE, suggesting that compensation consultants' specialized knowledge and expertise facilitate the use of RPE. Among all determinants, growth opportunities, common risk exposure, industry concentration, compensation consultants, and board size have a relatively large impact on RPE use.¹⁷

In summary, our evidence confirms the importance of common risk and CEOs' self-hedging ability in firms' decisions to use RPE, and it shows that firms with more independent boards and firms employing compensation consultants are more likely to use RPE. However, some of our results contradict prior evidence based on implicit tests (such as the effect of industry competition). We also find little evidence that the availability of similar peers affects firms' decision to use RPE. Taken together, the evidence suggests that firms consider both costs and benefits when deciding to use RPE in executive compensation contracts. A potential caveat is that firms could have adopted RPE contracts prior to the new disclosure regulation, and, hence, our causal inferences should be interpreted with caution.

We acknowledge that firms can use RPE implicitly through boards' subjective discretion, rather than pre-committing to a formulaic explicit RPE contract (Ferri 2009). We also note that the explanatory variables in Equation (2) likely capture factors that influence a firm's decision to use explicit versus implicit RPE. From the efficient-contracting perspective, when RPE can bring greater net benefits (e.g., higher common risk exposure), it is likely optimal for the principal and the agent to pre-commit to an explicit RPE contract. From the rent-seeking perspective, to the extent that more flexibility in the pay-setting process allows managers to reap greater private benefits, firms with weaker corporate governance (e.g., less independent boards) are less likely to commit to an explicit RPE contract. Furthermore, firms facing intense public scrutiny for transparent compensation disclosures (e.g., larger firms) are more likely to explicitly commit to RPE to cater to shareholder preference by reducing the extent of board discretion. We refrain, however, from drawing inferences on explicit versus implicit RPE use, as a thorough investigation of this topic is beyond the scope of our study.

¹⁷ Firms may use RPE to determine anywhere from a substantial portion to a minimal amount of executive pay. While it is difficult to ascertain the amount of executive pay that directly results from RPE use, we develop a rough proxy for the extent of RPE use (*RPE_prop*) based on the ratio of RPE-involved compensation (e.g., bonus and long-term incentive payouts, restricted stocks, or stock options) to total compensation for CEOs. We exclude RPE firms with *RPE_prop* less than 10 percent, and re-estimate Equation (2). The untabulated results are qualitatively similar as those reported. We also estimate Equation (2) using a Tobit regression, where we replace *RPE* with *RPE_prop*, and we obtain qualitatively similar results.

VI. SELECTION OF RPE PEERS

The overall benefits of an RPE contract depend critically on how efficiently the RPE peer group is selected.¹⁸ In this section, we first discuss determinants of RPE peer selection. Next, we compare characteristics between firms that are chosen to be in the RPE peer group and those that are not chosen. We then conduct a multivariate regression analysis of the peer-selection process.

Determinants of RPE Peer Selection

Common Risk Exposure

We measure potential peers' common risk exposure based on the similarity in operational strategies and industry environment, proxied by same industry membership (*Same_SIC2* and *Same_SIC3*), the similarity in the firms' scale and thus flexibility in responding to external shocks, proxied by S&P 1500 membership (*SP1500*) and S&P sub-index membership (*Same_SP*), and performance comovement measured by the correlation in stock returns between a potential peer and the RPE firm of interest ($CORR(PeerReturn, RPEReturn)$).

Similar Ability

To capture similarity in ability, we construct $|PeerMVE\_RPEMVE|$ and $|PeerReturn\_RPEReturn|$, defined as the absolute differences in market value of equity and annual stock returns, respectively, between a potential peer and the RPE firm of interest.

Self-Serving Bias

Since firms usually select peer groups used in compensation contracts at the beginning of the fiscal year (e.g., Cadman and Carter 2010; Faulkender and Yang 2010), we measure self-serving bias using potential peers' *expected* performance at the beginning of 2006. Specifically, we construct *PeerReturn_IndReturn*, defined as a potential peer's expected stock performance minus the median expected stock performance for the RPE firm's two-digit SIC industry (excluding the RPE firm of interest). We measure expected stock performance as analysts' first consensus 12-month target price forecast (issued within 90 days after the year beginning) minus the prevailing stock price, scaled by the prevailing stock price.

Symbolism

To test for symbolism considerations in the peer-selection process, we construct *PeerSales_IndSales* and *PeerMVE_IndMVE*, defined as the differences in sales and market value of equity, respectively, between a potential peer and the median levels for the RPE firm's two-digit SIC industry (excluding the RPE firm of interest).

Univariate Results

To get a preliminary assessment of which firms are chosen to be members of a RPE peer group, for each RPE firm that discloses self-selected peers, we identify all domestic firms listed on Compustat that are not selected as RPE peers.¹⁹ Recall that about 74 percent of RPE firms use stock returns as their performance metric when implementing RPE contracts (see Table 1, Panel C). These firms are likely to use stock returns as the basis for selecting peer firms (in addition to

¹⁸ We do not examine the peer-selection process of RPE firms that choose published market or industry indices as RPE peer groups (see Table 1, Panel E) because it is not feasible to test the selection of published indices due to difficulty in gathering the performance and composition information for a population of indices. We also do not examine firms' selection between published indices versus self-selected peer groups due to lack of related economic theories.

¹⁹ Some of our sample firms select foreign firms in forming their RPE peer groups. We do not include foreign firms listed on Compustat, because Compustat only includes foreign firms cross-listed in the United States and thus does not provide an unbiased sample for foreign firms. In addition, Compustat collects firms' reported accounting numbers, while cross-listed firms may report financial statements using non-U.S. accounting standards.

other considerations). To more accurately capture firms' peer-selection process, we further limit this analysis to the 135 RPE firms that employ self-selected peer groups and use stock returns as the RPE-based performance metric.²⁰

Table 5 provides descriptive statistics on selected peers and unselected peers. As shown in Panel A, selected peers are more likely to operate in the same industry as the RPE firms, and are more likely to belong to the S&P 1500 index and the same S&P sub-index as the RPE firms. Furthermore, performance comovement between potential peers and the RPE firms is much higher for selected peers than for unselected peers. These results suggest that selected peers share common shocks with the RPE firms to a greater extent than do unselected peers. Consistent with firms choosing peers with similar ability, we find that selected peers have closer size and stock performance to RPE firms than unselected peers. Panel A also reveals a potential peer-selection bias, in that selected peers have significantly lower expected stock performance than unselected peers. Finally, consistent with symbolism in setting CEO pay, we find that selected peers have larger sales and larger market capitalization than unselected peers.

Table 5, Panel B compares selected and unselected peers conditional on whether peers are in the RPE firm's industry. One intriguing finding is that *unselected* peers from RPE firms' industry (within-industry peers) display higher performance comovement than *selected* peers outside RPE firms' industry (outside-industry peers). Hence, cross-industry peer selection does not seem to improve performance comovement. Nevertheless, selected outside-industry peers have larger sales and market capitalizations than selected within-industry peers, consistent with RPE firms selecting larger firms outside their own industries for symbolism considerations.

Multivariate Regression Results

To further examine the selection of RPE peers, we estimate the following logistic regression for the 135 RPE firms that disclose self-selected peer groups and use stock returns as the RPE-based performance metric, with standard errors clustered by RPE firm:

$$\begin{aligned} Prob(RPE\_Peer_{ijt} = 1) = & \Phi(\alpha_0 + \alpha_1 Same\_SIC2_{ijt-1} + \alpha_2 Same\_SIC3_{ijt-1} + \alpha_3 SP1500_{ijt-1} \\ & + \alpha_4 Same\_SP_{ijt-1} + \alpha_5 CORR(PeerReturn, RPEReturn)_{ijt-1} \\ & + \alpha_6 |PeerMVE\_RPEMVE|_{ijt-1} + \alpha_7 |PeerReturn\_RPEReturn|_{ijt-1} \\ & + \alpha_8 PeerReturn\_IndReturn_{ijt-1} + \alpha_9 PeerSales\_IndSales_{ijt-1} \\ & + \alpha_{10} PeerMVE\_IndMVE_{ijt-1} + \varepsilon_{ijt}). \end{aligned} \quad (3)$$

In Equation (3), the dependent variable, RPE_Peer_{ijt} , is an indicator variable that equals 1 if a potential peer firm j (i.e., a domestic firm on Compustat) is chosen to be a member of the RPE peer group by RPE firm i for year t , and 0 otherwise.

Given that the relative proportion of selected versus unselected RPE peers is highly unbalanced, the estimated coefficients from logistic regressions could be significantly biased and inefficient (Owen 2007). To alleviate this concern, we follow the commonly used undersampling method to randomly remove unselected peers when estimating logistic regressions (e.g., Chawla et al. 2003). Specifically, for each RPE firm, we randomly select a set of peers from

²⁰ The untabulated results based on all RPE firms with self-selected RPE peers are qualitatively similar to those reported in Tables 5 and 6. We do not conduct similar tests on RPE firms using accounting performance metrics because the heterogeneity in RPE-based accounting performance metrics does not allow a sufficiently large sample size to examine any individual accounting performance metric. Nevertheless, if we use earnings as a general proxy for RPE-based accounting performance metrics, we obtain qualitatively similar inferences regarding the RPE peer selection for firms using accounting performance metrics in RPE contracts.

TABLE 5
Mean (Median) of Variables Used to Explain RPE Peer Selection for Selected RPE Peers and Unselected RPE Peers

Panel A: Selected versus Unselected RPE Peers

	Selected RPE Peers (n = 1,668)		Unselected RPE Peers (n = 363,618)		Unselected RPE Peers Used in Table 6 (n = 1,668)	
	Mean	Median	Mean	Median	Mean	Median
Common Risk Exposure						
<i>Same_SIC2</i>	0.634	1	0.034	0	0.032	0
<i>Same_SIC3</i>	0.454	0	0.016	0	0.016	0
<i>SP1500</i>	0.862	1	0.481	0	0.469	0
<i>Same_SP</i>	0.528	1	0.157	0	0.146	0
<i>CORR(PeerReturn, RPEReturn)</i>	0.364	0.474	0.197	0.265	0.186	0.250
Similar Ability						
<i> PeerMVE_RPEMVE </i>	12,081	3,222	15,326	4,005	13,738	3,939
<i> PeerReturn_RPEReturn </i>	0.244	0.175	0.389	0.294	0.366	0.274
Self-Serving Bias						
<i>PeerReturn_IndReturn</i>	-0.028	-0.029	0.057	-0.002	0.050	-0.002
Symbolism						
<i>PeerSales_IndSales</i>	8,533	2,911	2,303	152	2,223	125
<i>PeerMVE_IndMVE</i>	14,927	3,884	3,605	377	3,402	318

Panel B: Selected versus Unselected RPE Peers Conditional on RPE Firms' Industry

	Within RPE Firms' Industry				Out of RPE Firms' Industry			
	Selected RPE Peers (n = 1,058)		Unselected RPE Peers (n = 12,238)		Selected RPE Peers (n = 610)		Unselected RPE Peers (n = 351,380)	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median
Common Risk Exposure								
<i>Same_SIC2</i>	1	1	1	1	0	0	0	0
<i>Same_SIC3</i>	0.716	1	0.490	0	0	0	0	0

(continued on next page)

Panel B: Selected versus Unselected RPE Peers Conditional on RPE Firms' Industry

	Within RPE Firms' Industry				Out of RPE Firms' Industry			
	Selected RPE Peers (n = 1,058)		Unselected RPE Peers (n = 12,238)		Selected RPE Peers (n = 610)		Unselected RPE Peers (n = 351,380)	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median
<i>SP1500</i>	0.847	1	0.439	0	0.887	1	0.482	0
<i>Same_SP</i>	0.482	0	0.130	0	0.608	1	0.158	0
<i>CORR(PeerReturn, RPEReturn)</i>	0.412	0.550	0.346	0.452	0.280	0.314	0.192	0.258
Similar Ability								
<i> PeerMVE_RPEMVE </i>	8,651.800	2,537.870	17,367.370	5,204.560	16,193.140	4,863.120	15,255.730	3,970.910
<i> PeerReturn_RPEReturn </i>	0.217	0.153	0.301	0.204	0.294	0.231	0.392	0.298
Self-Serving Bias								
<i>PeerReturn_IndReturn</i>	-0.029	-0.029	0.047	-0.005	-0.027	-0.029	0.057	-0.002
Symbolism								
<i>PeerSales_IndSales</i>	5,834.930	1,986.020	1,925.470	146.780	10,247.360	4,984.220	1,918.150	101.423
<i>PeerMVE_IndMVE</i>	9,296.990	2,842.630	3,074.170	396.034	16,336.840	6,141.290	2,619.980	234.333

Selected RPE peers include domestic firms that are chosen by a RPE firm as RPE peers. Unselected RPE peers include all the other domestic firms that are not chosen by a RPE firm as RPE peers. Unselected RPE peers used in Table 6 include a set of randomly generated firms from Unselected RPE peers and are used in Table 6 regressions. We limit to RPE firms that use stock returns as the performance metric and self-selected peers in RPE contracts.

Variables are measured over or at the end of fiscal year 2005 (2006) if a firm's first proxy filed under the SEC's 2006 executive compensation disclosure rules is for fiscal year of 2006 (2007). *|PeerMVE_RPEMVE|* and *|PeerReturn_RPEReturn|* are winsorized at top one percentiles. *PeerReturn_IndReturn*, *PeerSales_IndSales*, and *PeerMVE_IndMVE* are winsorized at top and bottom one percentiles. The mean differences between Selected RPE peers and Unselected RPE peers are all statistically significant at less than 1 percent level based on two-tailed t-tests (untabulated).

Variable Definitions:

Same_SIC2 and *Same_SIC3* = 1 if the RPE firm and the potential peer share the same two-digit and three-digit SIC code, respectively, and 0 otherwise;
SP1500 = 1 if the potential peer is in the S&P 1500 index, and 0 otherwise;

Same_SP = 1 if the potential peer is in the same S&P 1500 sub-index (S&P 500, Mid-Cap 400, and Small-Cap 600) as the RPE firm, and 0 otherwise;

CORR(PeerReturn, RPEReturn) = Pearson correlation of annual stock returns between the RPE firm and the potential peer over the prior five years;

|PeerMVE_RPEMVE| and *|PeerReturn_RPEReturn|* = absolute values of the pair-wise differences between market value of equity and annual stock returns for the potential peer and market value of equity and annual stock returns for the RPE firm; and

PeerReturn_IndReturn, *PeerSales_IndSales*, and

PeerMVE_IndMVE = pair-wise differences between analysts' forecasted annual stock returns, sales, and market value of equity for the potential peer and median analysts' forecasted annual stock returns, sales, and market value of equity, respectively, for the RPE firm's industry (two-digit SIC code) excluding the RPE firm of interest (requires at least ten firms other than the RPE firm in the industry).

TABLE 6

**Logistic Regression of Peer Firm Selection for S&P 1500 Firms Using Self-Selected Peers
in Relative Performance Evaluation and Stock Returns as the Performance Metric**

Panel A: All RPE Peers

Independent Variables	Predicted Sign	Coefficient (p-value)	Change in Probability Q1 versus Q3 Values
Common Risk Exposure			
<i>Same_SIC2</i>	+	2.394 (<0.001)	0.416
<i>Same_SIC3</i>	+	1.005 (<0.001)	0.232
<i>SP1500</i>	+	0.525 (<0.001)	0.128
<i>Same_SP</i>	+	0.940 (<0.001)	0.219
<i>CORR(PeerReturn, RPEReturn)</i>	+	0.246 (0.001)	0.049
Similar Ability			
$ PeerMVE_RPEMVE $	—	—0.013 (<0.001)	—0.033
$ PeerReturn_RPEReturn $	—	—0.178 (0.197)	—0.014
Self-Serving Bias			
<i>PeerReturn_IndReturn</i>	—	—0.921 (<0.001)	—0.038
Symbolism			
<i>PeerSales_IndSales</i>	+	0.019 (0.016)	0.024
<i>PeerMVE_IndMVE</i>	+	0.017 (<0.001)	0.028
Intercept		—1.416 (<0.001)	
Number of selected/unselected peers		1,668/1,668	
Percent concordant/discordant		92.4/7.5	
Pseudo R^2		0.668	
Wald χ^2		895.429	
(p-value)		(<0.001)	

Panel B: Conditional on RPE Firms' Stock Performance

Independent Variables	Predicted Sign	Underperforming RPE Firms		Outperforming RPE Firms	
		Coefficient (p-value)	Change in Probability Q1 versus Q3 Values	Coefficient (p-value)	Change in Probability Q1 versus Q3 Values
Common Risk Exposure					
<i>Same_SIC2</i>	+	2.841 (<0.001)	0.445	2.068 (<0.001)	0.388

(continued on next page)

Panel B: Conditional on RPE Firms' Stock Performance

Independent Variables	Predicted Sign	Underperforming RPE Firms		Outperforming RPE Firms	
		Coefficient (p-value)	Change in Probability Q1 versus Q3 Values	Coefficient (p-value)	Change in Probability Q1 versus Q3 Values
<i>Same_SIC3</i>	+	0.633 (0.018)	0.153	1.362 (<0.001)	0.296
<i>SP1500</i>	+	0.213 (0.093)	0.053	1.094 (<0.001)	0.249
<i>Same_SP</i>	+	1.127 (<0.001)	0.255	0.777 (<0.001)	0.185
<i>CORR(PeerReturn, RPEReturn)</i>	+	0.089 (0.442)	0.016	0.267 (0.004)	0.057
Similar Ability					
<i> PeerMVE_RPEMVE </i>	—	—0.013 (0.001)	—0.037	—0.016 (0.069)	—0.034
<i> PeerReturn_RPEReturn </i>	—	0.072 (0.724)	0.005	—0.312 (0.132)	—0.026
Self-Serving Bias					
<i>PeerReturn_IndReturn</i>	—	—1.337 (<0.001)	—0.051	—0.588 (0.018)	—0.026
Symbolism					
<i>PeerSales_IndSales</i>	+	0.020 (0.163)	0.027	0.020 (0.012)	0.023
<i>PeerMVE_IndMVE</i>	+	0.014 (0.026)	0.028	0.019 (0.008)	0.029
Intercept		—1.303 (<0.001)		—2.058 (<0.001)	
Number of selected/ unselected peers		792/792		876/876	
Percent concordant/ discordant		93.1/6.7		92.1/7.8	
Pseudo R ²		0.690		0.665	
Wald χ^2 (p-value)		436.588 (<0.001)		446.964 (<0.001)	

Panel C: RPE Peers Not Used For Compensation Benchmarking Purposes versus Compensation-Benchmarking Peers Not Used for RPE Purpose

Independent Variables	RPE Peers Not Used for Pay Benchmarking		Benchmarking Peers Not Used for RPE	
	Coefficient (p-value)	Change in Probability Q1 versus Q3 Values	Coefficient (p-value)	Change in Probability Q1 versus Q3 Values
Common Risk Exposure				
<i>Same_SIC2</i>	2.566 (<0.001)	0.429	2.104 (<0.001)	0.391

(continued on next page)

Panel C: RPE Peers Not Used For Compensation Benchmarking Purposes versus Compensation-Benchmarking Peers Not Used for RPE Purpose

Independent Variables	RPE Peers Not Used for Pay Benchmarking		Benchmarking Peers Not Used for RPE	
	Coefficient (p-value)	Change in Probability Q1 versus Q3 Values	Coefficient (p-value)	Change in Probability Q1 versus Q3 Values
<i>Same_SIC3</i>	2.044 (<0.001)	0.385	0.545 (0.119)	0.133
<i>SP1500</i>	0.443 (0.043)	0.109	0.702 (<0.001)	0.169
<i>Same_SP</i>	0.852 (<0.001)	0.201	0.804 (<0.001)	0.191
<i>CORR(PeerReturn, RPEReturn)</i>	0.259 (0.068)	0.051	-0.521 (<0.001)	-0.103
Similar Ability:				
<i> PeerMVE_RPEMVE </i>	-0.021 (0.011)	-0.052	-0.009 (0.002)	-0.028
<i> PeerReturn_RPEReturn </i>	-0.306 (0.193)	-0.024	0.570 (0.004)	0.045
Self-Serving Bias				
<i>PeerReturn_IndReturn</i>	-1.236 (<0.001)	-0.049	1.338 (<0.001)	0.056
Symbolism				
<i>PeerSales_IndSales</i>	0.007 (0.610)	0.010	0.003 (0.688)	0.005
<i>PeerMVE_IndMVE</i>	0.022 (0.027)	0.036	0.013 (<0.001)	0.031
Intercept	-1.349 (<0.001)		-1.751 (<0.001)	
Number of selected/unselected peers	511/511		472/472	
Percent concordant/discordant	89.3/10.6		90.4/9.5	
Pseudo R^2	0.578		0.608	
Wald χ^2	228.344		229.045	
(p-value)	<0.001		<0.001	

The sample includes selected peers and unselected peers (generated from random sampling) for the 135 RPE firms choosing self-selected peer groups and using stock returns as the performance metric in relative performance evaluation. The dependent variable is *RPE_Peer*, which is equal to 1 if the firm (one Compustat domestic firm) is chosen as a peer for relative compensation evaluation purposes, and 0 otherwise. See Table 5 for the other variable definitions. "Underperforming (Outperforming) RPE Firms" include RPE firms with below-industry (above-industry) stock returns during the year prior to the SEC's new executive compensation disclosure rules. *|PeerMVE_RPEMVE|* and *|PeerReturn_RPEReturn|* are winsorized at the top one percentiles. *PeerReturn_IndReturn*, *PeerSales_IndSales*, and *PeerMVE_IndMVE* are winsorized at top and bottom one percentiles. For ease of exposition, coefficients on *PeerSales_IndSales*, *PeerMVE_IndMVE*, and *|PeerMVE_RPEMVE|* are multiplied by 1,000. The "Change in Probability" columns show the change in the probability of being selected as a RPE peer as the result of moving from the first to the third quartile value of the variable of interest (for indicator variables, moving from 0 to 1), holding all other variables constant at their mean values (for indicator variables, the benchmark probability is determined with the 0 value). Two-tailed p-values in parentheses are based on clustered standard errors at the RPE firm level. The reported p-values in parentheses are two-tailed.

domestic firms on Compustat that are not chosen by the RPE firm. For each RPE firm, we require

that the number of unselected peers in the random sample be equal to the number of selected peers. As reported in Table 5, Panel A, unselected peers generated from random sampling display similar characteristics to those of the population of unselected peers. To ensure the robustness of our results, we generate 100 random samples for each regression reported in Table 6, and find consistent results across each random generation.²¹

Table 6, Panel A reports results from estimating Equation (3). Consistent with the efficient-contracting view, firms that belong to the same industry as the RPE firm, firms included in the S&P 1500 index and from the same S&P sub-index as the RPE firm, and firms with higher return comovement are more likely to be chosen as RPE peers. We also find some evidence that similar firms are more likely to be selected as RPE peers—firms with closer size to the RPE firm are more likely to serve as RPE peers. With respect to selection bias, we find statistically significant evidence that firms with lower expected stock performance are more likely to be chosen as RPE peers. In line with the symbolism considerations, firms with larger sales and larger market capitalization than the RPE firm's industry median are more likely to be chosen as RPE peers.

Among all the factors affecting the peer-selection process, industry membership seems to have the greatest economic significance. For instance, firms from the same two-digit SIC industry as the RPE firm are about 42 percent more likely to be selected as RPE peers than firms are from different industries. S&P index membership also has substantial influence over the likelihood of being selected as a RPE peer. Performance comovement, size similarity, self-serving bias, and symbolism variables have smaller, but still significant, economic effects on RPE peer selection. While the peer-selection bias seems to have small economic significance, we are uncertain about the ultimate impact of peer-selection bias on executive compensation. Taken together, our findings support both efficient-contracting and rent-seeking behavior in the RPE peer-selection process.

We next examine the potential influence of RPE firms' performance on the relative importance of efficient contracting versus rent extraction in peer selection. We posit that underperforming managers are more concerned about job security and pay reduction, and, hence, have stronger incentives to manipulate the peer-selection process. To examine this issue, we re-estimate Equation (3) among RPE firms with below industry median versus above industry median stock returns during the year prior to the SEC's new disclosure regulation. The results are reported in Table 6, Panel B. Consistent with our conjecture, we find that the negative relation between potential peers' expected performance and the likelihood of being selected as a RPE peer is more (less) pronounced for RPE firms that underperform (outperform) their industries. Moreover, the positive impact of stock performance comovement over RPE peer selection is significant among outperforming RPE firms but insignificant among underperforming RPE firms. These findings suggest that superior performance leads efficient-contracting considerations to dominate in the RPE peer-selection process, probably due to managers worrying less about job security and pay reduction. In contrast, peer selection by underperforming firms appears to reflect rent-seeking behavior, probably because managers are concerned about reputation loss and pay reduction and thus have stronger incentives to manipulate the peer-selection process.²²

²¹ For instance, with respect to results reported in Table 6, Panel A, across the 100 random-sample regressions, the coefficients on *Same_SIC2*, *Same_SIC3*, *SP1500*, *Same_SP*, *lPeerMVE_RPEMVE*, *PeerReturn_IndReturn*, and *PeerMVE_IndMVE* are statistically significant as predicted for all 100 regressions. The coefficient on *PeerSales_IndSales* is statistically significant as predicted for 95 regressions. The coefficient on *CORR(PeerReturn, RPEReturn)* is statistically significant as predicted for 93 regressions. Finally, the coefficient on *lPeerReturn_RPEReturn* is statistically significant as predicted for 44 regressions.

²² Given that selected peers have larger sales and larger market capitalization than unselected peers, our findings may be confounded by size differences between selected and unselected peers. To mitigate this concern, we restrict the peer

RPE Peer Selection versus Compensation-Benchmarking Peer Selection

Many firms use compensation-benchmarking peers in setting the level of executive compensation, for the purpose of gauging an executive's reservation wage in the competitive labor market.²³ Recent studies suggest that firms select highly paid peers to justify excess CEO compensation (Bizjak et al. 2010; Faulkender and Yang 2010). RPE peers and benchmarking peers, however, not only differ in their theoretical justification, but also in the direction of peer-selection bias for peer performance: better performing firms tend to pay higher compensation and thus are more likely to be selected as compensation-benchmarking peers, whereas worse-performing firms are more likely to be selected as RPE peers to boost firms' relative performance and, hence, incentive compensation.

To shed some light on the peer selection for the purpose of RPE versus the purpose of pay benchmarking, we collect compensation-benchmarking peers for the 232 firms that use self-selected RPE peers in 2006. We find that 147 firms use two different self-selected peer groups²⁴ (one for RPE purposes and the other for pay-benchmarking purposes), with a mean (median) overlapping rate of 72 percent (81 percent), where the overlapping rate is defined as the number of common peers between the two peer groups divided by the size of the RPE peer group.

We then re-estimate Equation (3) by limiting the sample to selected RPE peers *not* included in the compensation-benchmarking peer group, together with randomly sampled peers. Similarly, we re-estimate Equation (3) by limiting the analysis to selected pay-benchmarking peers *not* used for RPE purposes, together with randomly sampled peers. To be consistent with analyses above, we further require firms to use stock returns as the RPE-based performance metric. Table 6, Panel C reports the regression results. For the selection of RPE peers not used for pay benchmarking, we find qualitatively similar results to those reported in Panel A. Interestingly, we observe that performance comovement negatively influences the selection of pay-benchmarking peers. Moreover, peers' expected stock performance has a positive impact on the likelihood of being chosen as a pay-benchmarking peer, consistent with firms favoring better-performing peers to boost executive pay. These results stand in contrast to the RPE peer-selection results, which support the notion that RPE peers and pay-benchmarking peers serve different purposes and their selection reflects different considerations.

Intertemporal Changes in RPE Peer Group Composition

To provide further evidence on RPE peer selection, we analyze intertemporal changes in RPE peer-group composition. If peer selection conforms to efficient contracting, then we expect added peers to exhibit greater ability to remove common risk and greater similarity with the RPE firm

population to S&P 1500 firms, and re-estimate Equation (3). This sample restriction ignores the fact that about 30 percent of selected RPE peers do not belong to the S&P 1500 index (see Table 1, Panel F) which potentially lowers the test power. Nevertheless, we find qualitatively similar results as those reported in Table 6, Panel A (untabulated). We also find qualitatively similar results as those reported in Table 6, Panel B. In particular, the coefficient on *PeerReturn_IndReturn* is significantly negative for underperforming RPE firms, but insignificantly positive for outperforming RPE firms (untabulated).

²³ Consistent with the efficient use of benchmarking peers, recent studies show that benchmarking peers represent a set of companies against which a firm competes for executive talent (Bizjak et al. 2008, 2010; Cadman and Carter 2010; Albuquerque et al. 2009).

²⁴ Starting with 232 firms, we are able to find 228 firms' pay-benchmarking peer groups from reading CD&A reports. For these 228 firms, we find that 20 firms use either a broader definition of peers (such as an index) or proprietary compensation survey data for pay-benchmarking purposes. For the 208 firms that use self-selected pay-benchmarking peers, the mean (median) size of pay-benchmarking peer groups is 17 (15) peers. Of these 208 firms, 36 firms use the same peer groups for both pay-benchmarking and RPE purposes; 25 firms have RPE peer groups that are a subset of their pay-benchmarking peer groups; 3 firms use completely different peer groups for two purposes; and the remaining 144 firms have some overlap between the two peer groups.

than dropped peers. If, on the other hand, firms choose poorly performing peers when forming RPE peer groups, then we expect added peers to exhibit worse performance than dropped peers.

To identify the sample, we start with the 232 firms that disclose self-selected RPE peers in 2006. We find that 180 firms continue using RPE in 2007, among which 155 firms disclose self-selected RPE peer groups. Of these 155 firms, 22 firms use the same RPE peer groups and 133 firms modify their RPE peer groups in 2007, with 74 percent of 2006 RPE peers retained in 2007. The final sample includes 75 firms that both dropped and added RPE peers in 2007 and have available data on required variables.

Table 7 reports differences in the factors affecting the selection of RPE peers between added and dropped peers. We find that added peers are closer in size to RPE firms than dropped peers, consistent with more (less) similar firms being added to (dropped from) the RPE peer group. Moreover, added peers' industry-adjusted expected performance is significantly negative, whereas dropped peers' industry-adjusted expected performance is insignificant. While this finding supports self-serving bias in RPE peer selection, the contrast between added and dropped peers is statistically insignificant. This weak evidence could reflect low test power due to stable peer-group composition and/or increased public scrutiny of the peer-selection process after the SEC's new regulation prompting firms to restrain opportunistic behavior in peer selection.

VII. CONCLUSION

We examine the explicit use of relative performance evaluation (RPE) and related peer groups based on S&P 1500 firms' first proxy disclosures under the SEC's 2006 executive compensation

TABLE 7
Mean (Median) of Variables Used to Explain RPE Peer Selection for Peers Added, Dropped, and Unchanged in 2007

	Peers Added (1)	Peers Dropped (2)	Peers Unchanged (3)	Paired t-statistic (1)-(2)
Common Risk Exposure				
<i>Same_SIC2</i>	0.614	0.653	0.703	-1.63
<i>Same_SIC3</i>	0.481	0.465	0.552	1.27
<i>SP1500</i>	0.720	0.707	0.789	0.15
<i>Same_SP</i>	0.590	0.507	0.643	1.49
<i>CORR(PeerReturn, RPEReturn)</i>	0.357	0.424	0.394	-1.54
Similar Ability				
<i> PeerMVE_RPEMVE </i>	15,428.350	21,772.750	15,703.120	-1.68*
<i> PeerReturn_RPEReturn </i>	0.232	0.253	0.222	-0.89
Self-Serving Bias				
<i>PeerReturn_IndReturn</i>	-0.035	-0.014	-0.038	-1.55
Symbolism				
<i>PeerSales_IndSales</i>	10,252.390	14,182.150	11,710.980	-0.70
<i>PeerMVE_IndMVE</i>	18,447.670	22,862.210	19,342.390	-0.52

* Indicates significance at less than the 10 percent level based on two-tailed paired t-tests of the mean differences.

Bold figures indicate significance level at less than 5 percent.

The sample includes 75 firms that report two different RPE peer groups for 2006 and 2007 and have available data on required variables. All variables are measured at the beginning of 2007. *PeerReturn* is peer firms' expected stock performance (based on analysts' target price forecasts at the beginning of 2007). See Table 5 for the other variable definitions.

disclosure rules. We demonstrate that a lack of knowledge of both RPE peer-group composition and the link between RPE-based performance targets and future peer-performance cloud inferences drawn from implicit tests. These findings highlight the limitations of the implicit approach and underscore the importance of incorporating explicit RPE contract details in testing for RPE use. We also find that firms consider both costs and benefits of RPE as an incentive mechanism when deciding to use RPE. Analyses of self-selected RPE peers lend support to the efficient use of RPE as well as a self-serving bias in forming RPE peer groups.²⁵ In particular, RPE firms' superior performance leads efficient-contracting considerations to drive the selection of RPE peers, whereas underperforming RPE firms exhibit greater rent-seeking behavior in the RPE peer-selection process.

One limitation of our analyses is that we are unable to identify firms that use RPE implicitly without pre-committing to explicit RPE contracts. Future research could examine both explicit and implicit contracting to better understand RPE as an incentive device. Furthermore, while our study sheds light on target-setting in executive compensation contracts, our analysis likely omits other factors that boards consider in setting performance targets for executives. The peer-selection bias we document, for example, could reflect boards' conscious efforts to balance various factors in setting performance targets (Merchant and Manzoni 1989). Examining interactions among various factors would be an interesting avenue for future research.

APPENDIX A TIMELINE OF INTRODUCING THE SEC'S NEW EXECUTIVE COMPENSATION DISCLOSURE RULES

October 16, 1992 to January 27, 2006: *Executive Compensation Disclosure*

The 1992 regulation eschewed a mostly narrative disclosure approach adopted in 1983 in favor of formatted tables that capture all executives' compensation.

January 27, 2006: SEC released the proposed rule *Executive Compensation and Related Party Disclosure*

The proposed rule combines a broader-based tabular presentation with improved narrative disclosure supplementing the tables.

January 27, 2006 to April 10, 2006: Open for public comments

SEC had received more than 20,000 comment letters, most of which supported the new disclosure reform.

August 29, 2006: SEC released the final rule *Executive Compensation Disclosure*

August 29, 2006 to October 23, 2006: Open for public comments

Effective Date:

Companies must comply with these disclosure requirements in Forms 8-K for triggering events that occur on or after November 7, 2006 and in Forms 10-K and 10-KSB for fiscal years ending on or after December 15, 2006. Companies other than registered investment companies must comply

²⁵ Other than RPE peer selection, there exist alternative channels through which firms can influence the achievability of RPE-based performance targets and, hence, executive pay (such as through raising or lowering the hurdle rate of RPE-based performance targets or selecting different RPE-based performance metrics). We acknowledge that our inferences concerning the efficiency or opportunism in the RPE peer selection do not necessarily imply optimal or excessive executive pay, because alternative design features of RPE contracts (such as the hurdle rate or performance metrics) might either introduce greater rent extraction or offset the selection bias in the RPE peer-selection process.

with these disclosure requirements in Securities Act registration statements and Exchange Act registration statements (including pre-effective and post-effective amendments), and in any proxy or information statements filed on or after December 15, 2006 that are required to include Item 402 and 404 disclosure for fiscal years ending on or after December 15, 2006. Registered investment companies must comply with these disclosure requirements in initial registration statements and post-effective amendments that are annual updates to effective registration statements on Forms N-1A, N-2 (except those filed by business development companies) and N-3, and in any new proxy or information statements, filed with the Commission on or after December 15, 2006. (directly quoted from the SEC)

APPENDIX B

CODING CRITERIA AND EXAMPLES OF PROXY DISCLOSURES ABOUT RELATIVE PERFORMANCE EVALUATION (RPE)

Identifying RPE Firms from Compensation Discussion and Analysis (CD&A) Reports

If a firm states that at least one component of executive compensation (e.g., annual bonus, restricted stock, stock option, long-term incentive payout) is determined based on firm performance relative to a group of peers, then the firm is identified as a RPE firm; otherwise, the firm is identified as a non-RPE firm. To be conservative, we classify firms that comment favorably on the principle of RPE, but do not disclose details on the use of RPE in setting compensation, as non-RPE firms.

Since the use of compensation-benchmarking peer groups is much more prevalent than the use of RPE peer groups in S&P 1500 firms, it is important to distinguish firms using compensation benchmarking from firms using RPE. The SEC defines compensation benchmarking as “benchmarking of total compensation, or any material element of compensation” (SEC Final Rule 33-8732a, Item 402(b)(2)(xiv)). Consistent with this definition, we classify firms using peer groups to benchmark the *level* of executive compensation as pay-benchmarking firms. Consequently, if our reading of the CD&A report for a sample firm suggests that the firm uses peer groups to solely benchmark the *level* of pay without using peer groups to determine the *performance target* when evaluating the firm’s performance, then the firm is coded as a non-RPE firm.

Below are excerpts from the CD&A reports in the first proxy statements filed after December 16, 2006 for four companies. The first three companies are classified as RPE firms. The last company is classified as a non-RPE firm due to lack of details about RPE use in compensation contracts. Discussions about RPE use are bolded.

General Electric Co

Performance share units (PSUs). Since 2003, we have compensated the CEO with PSUs in lieu of any other equity incentive compensation... The receipt of shares underlying PSUs is determined entirely by the performance of the company against two key metrics: an internal metric that measures the cash-producing capability of the company and **an external metric that measures the performance of the company against a broad market index...** PSUs will convert into shares of GE stock at the end of the five-year performance period only if the specified performance objectives have been achieved. Half of the PSUs will convert into shares of GE stock only if GE’s cash flow from operating activities, adjusted to exclude the effect of unusual events, has grown an average of 10 percent or more per year over the five-year performance period. Otherwise, they will be cancelled. **The remaining PSUs will convert into shares of GE stock only if GE’s total shareowner return meets or exceeds that of the S&P 500 over the five-year performance period. Otherwise, they will be cancelled.**

PG&E Corp

In establishing levels of executive compensation, each year the Committee reviews the appropriateness of the comparator groups used to assess the competitiveness of PG&E Corporation's compensation programs (Pay Comparator Group) and PG&E Corporation's corporate performance (**Performance Comparator Group**)... The primary comparator group used for purposes of setting 2006 officer compensation consists of all companies listed in the Dow Jones Utility Index and the Standard & Poor's Electrics Index, and all California investor-owned utilities (the "Pay Comparator Group")... This group of companies is broad enough to provide statistical validity and data availability, represents the segment of the market where PG&E Corporation and Pacific Gas and Electric Company recruit officers with industry-specific experience, and is determined on an objective and transparent basis. **For purposes of corporate performance comparisons (including the relative total shareholder return measured for the 2006–2008 performance share award cycle), the Committee uses a subgroup of 12 companies that have similar characteristics and business models as PG&E Corporation (the "Performance Comparator Group"):** Ameren Corporation, American Electric Power, CenterPoint Energy, Inc., Consolidated Edison, Entergy Corporation, FPL Group, NiSource Inc., Pinnacle West Capital, Progress Energy, Inc., Southern Company, TECO Energy, and Xcel Energy. This group of companies is a subset of the Pay Comparator Group and, like PG&E Corporation, is focused on core regulated-utility activities with either a distribution or an integrated-utility focus.

Performance Shares

The payment for performance shares will be in cash and will be calculated by multiplying (1) the number of vested performance shares, (2) the average closing price of PG&E Corporation common stock over the last 30 calendar days of the year preceding the vesting date, and (3) a payout factor based on corporate performance... **There will be no payout for TSR ("total shareholder return") performance below the 25th percentile of the Performance Comparator Group; there will be a 25 percent payout if TSR is at the 25th percentile; there will be a 100 percent payout if TSR is at the 75th percentile; and there will be a 200 percent payout if PG&E Corporation's TSR ranks first in the Performance Comparator Group. If PG&E Corporation's TSR is between the 25th percentile and the 75th percentile, or above the 75th percentile, award payouts will be determined by straight-line interpolation, adjusted to round numbers (i.e., the nearest multiple of five).**

Cleco Corp

Our Compensation Committee uses two comparator groups to design executive-officer compensation plans and track comparable performance of those plans. These groups are referred to as comparator group(s), peer group(s), peers or the competitive market throughout this discussion. The Base Peer Group was selected based on the companies being of approximate size and scope to Cleco (after regression analysis for size differences), employing similar labor and talent pools and having their executive officer compensation data being available to the outside independent consultant who analyzes the market data for the Compensation Committee. The Compensation Committee considers the availability of such detailed market data to be critical in making comparative compensation decisions... **The Incentive Peer Group was selected by the Compensation Committee in order to measure the actual performance results of our incentive plans. The Incentive Peer Group is based on the companies being part of a recognized stock market index, as well as being in the same general industry classification system. This helps ensure the Compensation Committee evaluates our actual incentive plan performance against a group of companies whose scope of operations and market capitalization is similar to Cleco's. Data from the peer groups are an integral part of the decision process used by the Compensation Committee in determining the design, component parts and levels of awards contained in our executive officer pay programs.**

Base Peer Group

For 2006 and 2007, executive officer compensation levels were evaluated using the Base Peer Group. This includes base salary, annual and long-term incentive plan targets, other potential equity awards and total compensation. The Base Peer Group consists of companies that are generally either in the Edison Electric Institute ("EEI") Index or the S&P Small and MidCap Electric Utilities Index. We are included in both indices.

Incentive Peer Group

For 2006, the relative actual performance of the financial measures used in our annual and long-term incentive plans was determined using the Incentive Peer Group. **The Incentive Peer Group consisted of the following 15 companies contained in the S&P Small and MidCap Electric Utilities Index: Allete, Inc.; Central Vermont Public Service; DPL, Inc.; Duquesne Light Holdings, Inc.; El Paso Electric Company; Great Plains Energy Inc.; Green Mountain Power Corporation; Hawaiian Electric Industries; IDACORP, Inc.; Northeast Utilities; Pepco Holdings, Inc.; Sierra Pacific Resources; UIL Holdings Corporation; UniSource Energy Corporation; and Westar Energy, Inc.** The same 15 companies will be used as the Incentive Peer Group in 2007.

Consolidated Graphics Inc

Overall Compensation Philosophy and Policies

Our compensation philosophy regarding members of the Board and the Company's Executive Officers is to maintain compensation policies which align compensation with the Company's overall business strategy, values and management initiatives. The policies are intended to (1) reward individuals for long-term strategic management and enhancement of shareholder value; (2) **support a performance-oriented environment that rewards achievement of internal Company goals and recognizes the Company's performance compared to the performance of similarly situated companies**; (3) attract and retain individuals whose abilities are considered essential to the long-term future and competitiveness of the Company; and (4) align the financial interests of the Company's directors and Executive Officers with those of the shareholders.

REFERENCES

- Aggarwal, R., and A. Samwick. 1999. Executive compensation, strategic competition and relative performance evaluation: Theory and evidence. *The Journal of Finance* 54 (6): 1999–2043.
- Albuquerque, A. 2009. Peer firms in relative performance evaluation. *Journal of Accounting and Economics* 48 (1): 69–89.
- . 2007. Do growth option firms use less relative performance evaluation? Working paper, Boston University.
- , G. De Franco, and R. Verdi. 2009. Peer choice in CEO compensation. Working paper, Boston University.
- Antle, R., and A. Smith. 1986. An empirical investigation of the relative performance evaluation of corporate executives. *Journal of Accounting Research* 24 (1): 1–39.
- Bannister, J., and H. Newman. 2003. Analysis of corporate disclosures on relative performance evaluation. *Accounting Horizons* 17 (3): 235–246.
- Barro, J., and R. Barro. 1990. Pay, performance, and turnover of bank CEOs. *Journal of Labor Economics* 8 (4): 448–481.
- Bebchuk, L., and J. Fried. 2003. Executive compensation as an agency problem. *The Journal of Economic Perspectives* 17 (3): 71–92.
- Bertrand, M., and S. Mullainathan. 2001. Are CEOs rewarded for luck? The ones without principals are. *The Quarterly Journal of Economics* 116 (3): 901–932.
- Bizjak, J., M. Lemmon, and L. Naveen. 2008. Does the use of peer groups contribute to higher pay and less

- efficient compensation? *Journal of Financial Economics* 90 (2): 152–168.
- , ———, and T. Nguyen. 2010. Are all CEOs above average? An empirical analysis of compensation peer groups and pay design. *Journal of Financial Economics* (forthcoming).
- Bol, J., T. Keune, E. M. Matsumura, and J. Y. Shin. 2010. Supervisor discretion in target setting: An empirical investigation. *The Accounting Review* 85 (6): 1861–1886.
- Bushman, R., Q. Chen, E. Engel, and A. Smith. 2004. Financial accounting information, organizational complexity and corporate governance systems. *Journal of Accounting and Economics* 37 (2): 167–201.
- Byrd, J., M. Johnson, and S. Porter. 1998. Discretion in financial reporting: The voluntary disclosure of compensation peer groups in proxy statement performance graphs. *Contemporary Accounting Research* 15 (1): 25–52.
- Cadman, B., and M. Carter. 2010. Compensation peer groups and their relation with CEO compensation. Working paper, The University of Utah.
- , ———, and S. Hillegeist. 2010. The incentives of compensation consultants and CEO pay. *Journal of Accounting and Economics* 49 (3): 263–280.
- Carter, M., C. Ittner, and S. Zechman. 2010. Explicit relative performance evaluation in performance-vested equity grants. *Review of Accounting Studies* 14 (2–3): 269–306.
- Chawla, N. V., N. Japkowicz, and A. Kolcz. 2003. *Learning from Imbalanced Data Sets*. Proceedings of the ICML 2003 Workshop, Ottawa, Canada.
- Cremers, K., and V. Nair. 2005. Governance mechanism and equity prices. *The Journal of Finance* 60 (6): 2859–2894.
- DeFond, M., and C. Park. 1999. The effect of competition on CEO turnover. *Journal of Accounting and Economics* 27 (1): 35–56.
- Dikolli, S., C. Hofmann, and T. Pfeiffer. 2010. Peer performance aggregation and relative performance evaluation. Working paper, Duke University.
- Dye, R. 1984. The trouble with tournaments. *Economic Inquiry* 22 (1): 147–149.
- Faulkender, M., and J. Yang. 2010. Inside the black box: The role and composition of compensation peer groups. *Journal of Financial Economics* 96 (2): 257–270.
- Ferri, F. 2009. Discussion of: Explicit relative performance evaluation in performance-vested equity grants. *Review of Accounting Studies* 14 (2–3): 307–313.
- Garvey, G., and T. Milbourn. 2003. Incentive compensation when executives can hedge the market: Evidence of relative performance evaluation in the cross section. *The Journal of Finance* 58 (4): 1557–1582.
- , and T. Milbourn. 2006. Asymmetric benchmarking in compensation: Executives are rewarded for good luck but not penalized for bad. *Journal of Financial Economics* 82 (1): 197–226.
- Gerakos, J., C. Ittner, and D. Larcker. 2007. The structure of performance-vested stock option grants. In *Essays on Accounting Theory in Honor of Joel S. Demski*, edited by R. Antle, P. Liang, and F. Gjesdahl. New York, NY: Springer.
- Gibbons, R., and K. Murphy. 1990. Relative performance evaluation for chief executive officers. *Industrial & Labor Relations Review* 43 (3): 30–51.
- Green, J., and N. Stokey. 1983. A comparison of tournaments and contracts. *The Journal of Political Economy* 91 (3): 349–364.
- Hannan, L., R. Krishnan, and D. Newman. 2008. The effects of disseminating relative performance feedback in tournament and individual performance compensation plans. *The Accounting Review* 83 (4): 893–913.
- Himmelberg, C., and R. Hubbard. 2000. Incentive pay and market for CEOs: An analysis of pay-for-performance sensitivity. Working paper, Columbia University.
- Holmstrom, B. 1982. Moral hazard in teams. *The Bell Journal of Economics* 13 (2): 324–340.
- Hvide, H. 2002. Tournament rewards and risk taking. *Journal of Labor Economics* 20 (4): 877–898.
- Janakiraman, S., R. Lambert, and D. Larcker. 1992. An empirical investigation of the relative performance evaluation hypothesis. *Journal of Accounting Research* 30 (1): 53–69.
- Joh, S. 1999. Strategic managerial incentive compensation in Japan: Relative performance evaluation and product market collusion. *The Review of Economics and Statistics* 81 (2): 303–313.
- Knoeber, C., and W. Thurman. 1994. Testing the theory of tournaments: An empirical analysis of broiler

- production. *Journal of Labor Economics* 12 (2): 155–179.
- Lazear, E., and S. Rosen. 1981. Rank-order tournaments as optimum labor contracts. *The Journal of Political Economy* 89 (5): 841–864.
- Leone, A., and S. Rock. 2002. Empirical tests of budget ratcheting and its effect on managers' discretionary accrual choices. *Journal of Accounting and Economics* 33 (1): 43–67.
- Matsumura, E. M., and J. Y. Shin. 2006. An empirical analysis of an incentive plan with relative performance measures: Evidence from a postal service. *The Accounting Review* 81 (3): 533–566.
- Merchant, K. A., and J.-F. Manzoni. 1989. The achievability of budget targets in profit centers: A field study. *The Accounting Review* 64 (3): 539–558.
- Murphy, K. 1999. Executive compensation. In *Handbook of Labor Economics*, Vol. 3, edited by Ashenfleter, O., and D. Card. Amsterdam, The Netherlands: North-Holland.
- . 2000. Performance standards in incentive contracts. *Journal of Accounting and Economics* 30 (3): 245–278.
- , and T. Sandino. 2010. Executive pay and “independent” compensation consultants. *Journal of Accounting and Economics* 49 (3): 247–262.
- Nalebuff, B., and J. Stiglitz. 1983. Prizes and incentives: Towards a general theory of compensation and competition. *The Bell Journal of Economics* 14 (1): 21–43.
- O'Keeffe, M., W. Viscusi, and R. Zeckhauser. 1984. Economic contests: Comparative reward schemes. *Journal of Labor Economics* 2 (1): 27–56.
- Owen, A. 2007. Infinitely imbalanced logistic regression. *Journal of Machine Learning Research* 8 (4): 761–773.
- Prendergast, C. 1999. The provision of incentives in firms. *Journal of Economic Literature* 37 (1): 7–63.
- Rajgopal, S., T. Shevlin, and V. Zamora. 2006. CEO's outside employment opportunities and the lack of relative performance evaluation in compensation contracts. *The Journal of Finance* 61 (4): 1813–1844.
- Rose, N., and A. Shepard. 1997. Firm diversification and CEO compensation: Managerial ability or executive entrenchment? *The Rand Journal of Economics* 28 (3): 489–514.
- Rosen, S. 1986. Prizes and incentives in elimination tournaments. *The American Economic Review* 76 (4): 701–715.
- Westphal, J., and E. Zajac. 1994. Substance and symbolism in CEO's long-term incentive plans. *Administrative Science Quarterly* 39 (3): 267–390.
- Zajac, E., and J. Westphal. 1995. Accounting for the explanations of CEO compensation: substance and symbolism. *Administrative Science Quarterly* 40 (2): 283–308.

Inflation and Nominal Financial Reporting: Implications for Performance and Stock Prices

Yaniv Konchitchki
University of Southern California

ABSTRACT: The monetary unit assumption of financial accounting assumes a stable currency (i.e., constant purchasing power over time). Yet, even during periods of low inflation or deflation, nominal financial statements violate this assumption. I posit that, while the effects of inflation are not recognized in nominal statements, such effects may have economic consequences. I find that unrecognized inflation gains and losses help predict future cash flows as these gains and losses turn into cash flows over time. I also find significant abnormal returns to inflation-based trading strategies, suggesting that stock prices do not fully reflect the implications of the inflation effects for future cash flows. Additional analysis reveals that stock prices act as if investors do not fully distinguish monetary and nonmonetary assets, which is fundamental to determining the effects of inflation. Overall, this study is the first to show that, although inflation effects are not recognized in nominal financial statements, they have significant economic consequences, even during a period in which inflation is relatively low.

Keywords: *inflation; asset pricing; information; financial reporting; abnormal returns; cash flows; capital markets.*

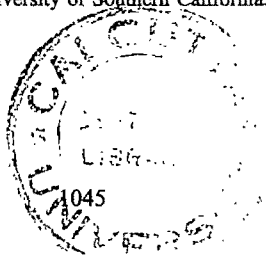
Data Availability: *Data are available from public sources indicated in the text.*

I. INTRODUCTION

The U.S. financial reporting regime is nominal, which assumes no changes to the purchasing power of the dollar over time. That is, the accounting amounts reported in financial statements based on U.S. Generally Accepted Accounting Principles (GAAP) are called nomi-

I am indebted to my dissertation committee members, Mary Barth (chair), Bill Beaver, and Ron Kasznik, for valuable guidance on this study, and to Steven Kachelmeier (senior editor), Mark Trombley (editor), and two anonymous referees for invaluable comments and suggestions. I thank Nissim Aranya, Snehal Banerjee, Scott Joslin, and Gil Sadka for helpful discussions. I also thank Carlos Corona, Mark DeFond, Dan Givoly, Ilan Guttman, Wayne Landsman, Charles Lee, Scott Richardson, Ronnie Sadka, Ron Siegel, Richard Sloan, Mark Soliman, Steve Stubben, K.R. Subramanyam, Dan Taylor, and seminar participants at Barclays Global Investors, Columbia University, Georgetown University, Rice University, Stanford University, University of Miami, University of Minnesota, The University of North Carolina at Chapel Hill, University of Southern California, and the University of Toronto for helpful comments and suggestions. I am thankful for financial support from Stanford University and University of Southern California.

Editor's note: Accepted by Mark Trombley.



Submitted: February 2009
Accepted: September 2010
Published Online: April 2011

nal because they are not adjusted for inflation.¹ Yet, the annual U.S. inflation rate over the past two decades has averaged to 3 percent. When purchasing power is not constant, a nominal reporting system that combines monetary amounts from different periods violates the Monetary Unit accounting assumption of a stable currency (SFAC 5, Financial Accounting Standards Board [FASB] 1984) and decreases comparability both across firms and over time. Although the effects of inflation on financial reporting received considerable attention during the late 1970s when inflation was relatively high, inflationary effects have been mostly ignored in the more modest inflationary environment since that time.

Nominal financial statements, by their nature, do not account for gains and losses attributable to changes in purchasing power over time. For example, whereas the erosion of a firm's monetary assets (e.g., cash) attributable to inflation is a loss to the firm, the erosion of a firm's monetary liabilities (e.g., debt) is a gain. Further, whereas inflation-adjusted amounts of nonmonetary items (e.g., land) accumulate inflationary effects over time to reflect changes in purchasing power, such effects are not recognized in nominal financial statements. The difference between inflation-adjusted and nominal earnings represents unrecognized gains and losses from inflation. In this study, I investigate the economic consequences of omitting these inflation effects from nominal financial statements. Specifically, I test whether unrecognized gains and losses from inflation help predict future cash flows, and whether investors incorporate this information into their investment decisions.

This study is most closely related to inflation accounting studies that investigate the implications of inflation in the U.S. during the 1970s and 1980s (e.g., Beaver 1979; Easman et al. 1979; Beaver et al. 1980; Gheyara and Boatsman 1980; Ro 1980; Watts and Zimmerman 1980; Beaver and Landsman 1983). These studies address questions primarily related to whether the effects of inflation are associated with contemporaneous annual and short-window stock returns. A conclusion from these studies is that inflation-adjusted (or, interchangeably, IA) data are inconsequential for making financial decisions.²

I take a different approach by considering the possibility that inflation can have implications over a period longer than the contemporaneous year. In particular, if unrecognized inflation gains and losses (or simply *inflation gains*, for brevity) are realized over time, then they are likely to help predict firms' future cash flows.³ Moreover, if the stock market does not fully account for such implications for future cash flows, then inflation gains can be associated with future stock returns. Thus, by focusing attention on possible longer-horizon inflation effects, I shed new light on the extent to which the stock market incorporates inflation information, and on the pricing

¹ In the text I refer to inflation because deflation has rarely occurred in the U.S.; however, the study pertains to both inflation and deflation because deflation can be viewed as negative inflation.

² This research mainly investigates questions related to the adoption of the Securities and Exchange Commission (SEC)'s Accounting Series Release (ASR) 190 in 1976 and Statement of Financial Accounting Standard (SFAS) 33 in 1979, which are no longer effective, that were mandated in response to high inflation during the period. Another related area follows Modigliani and Cohn (1979), who hypothesize that the stock market is inefficient because it suffers from inflation illusion (e.g., Ritter and Warr 2002; Campbell and Vuolteenaho 2004; Basu et al. 2006; Chordia and Shivakumar 2005). Another area attempts to explain the association between stock returns and unexpected inflation (e.g., Summers 1981; Bernard 1986). Whereas prior studies in this area examine how raw returns vary with respect to unexpected inflation, I investigate how investors process inflation-adjusted information and whether *ex ante* portfolios can be formed to generate future abnormal returns. I also extend prior work by developing a measure that captures the total information content of inflation that is omitted when reporting nominal amounts (e.g., purchasing dates for different asset classes, as explained in the Appendix), and in contrast to prior studies, which analyze the effects of unexpected inflation on realized returns, I analyze the effects of expected inflation on expected returns. Other related areas investigate inflation effects in a variety of contexts, including the bond market, non-U.S. countries with high inflation, employment determination, dividend decision, and inflation tax (e.g., Phelps 1967; Fama and Schwert 1977; Bar-Yosef and Lev 1983; Gordon 2001; Kothari and Shanken 2004; Davis-Friday and Gordon 2005).

³ I use the terms "inflation gains" and "unrecognized inflation gains" interchangeably throughout the study.

implications of inflation-adjusted data. I also extend prior research by providing insight into how investors process inflation-adjusted information, thereby providing evidence on the mechanism through which inflation impacts future investment decisions.

The first question I address is whether inflation gains can help predict firms' future cash flows. The inflation effects can translate to future cash flows because higher unrecognized inflation gains accumulated in nonmonetary assets should result in higher cash flows from operations when the assets are used (in the case of property, plant, and equipment [PPE]) or sold (in the case of inventory) through many types of business activities, leading to a positive association between unrecognized inflation gains and future cash flows.

I address this question using two tests. In the first test, I investigate how inflation distorts nominal amounts. To do so, I develop an algorithm that extracts inflation-adjusted data from firms' nominal financial statements. More specifically, the algorithm adjusts nominal financial statements for inflation, firm-by-firm, using the distinction between monetary and nonmonetary amounts, the life cycles of assets and liabilities, and the clean surplus relation. An important feature of the algorithm is that it allows inflation effects to vary across firms and over time, capturing the fact that inflation affects firms differently depending on the structure of their assets and liabilities. Further, the algorithm and the inflation-adjusted amounts in this study rely on the historical cost measurement attribute, which has reliability and objectivity advantages. This is the same system of accounting that underlies the current nominal reporting regime, except that amounts are stated in common units. This test evidences considerable cross-sectional and time-series variation in inflation effects. To verify the external validity of the algorithm, I implement it on a sample of Israeli firms for which I hand-collect inflation-adjusted and nominal data—Israeli firms were required to disclose both sets of figures until 2003. This verification analysis (see the Appendix) shows that the algorithm provides reasonable estimates of inflation-adjusted financial statement amounts.

In the second test, I examine whether unrecognized inflation gains are realized over time, and thus can help to predict future performance. In particular, I test whether unrecognized inflation gains improve forecasts of cash flows from operations for horizons from one to four years relative to a benchmark model developed in Barth et al. (2001). I find that unrecognized inflation gains are associated with future cash flows from operating activities in each of the four subsequent years.

Next, I focus on the stock market effects of inflation. If inflation-adjusted information has implications for future cash flows, then investors are likely to incorporate it into their investment decisions. However, using inflation-adjusted data can be costly because inflation-adjusted data are not reported and processing such data is more complicated than processing nominal data (Beaver and Landsman 1983). Because mispricing can arise when information is costly to obtain and process (Grossman and Stiglitz 1980; Ball 1994), even under the semi-strong form of market efficiency, this raises the question of whether investors fully incorporate the implications of inflation information for future cash flows in their equity-pricing decisions. I therefore investigate whether unrecognized inflation gains can be used to generate significant abnormal returns.

Specifically, given my findings that inflation gains have predictive power for future cash flows, future returns are likely to depend on stock market expectations regarding these gains. If the market fully incorporates information about inflation gains, then stock prices will correctly reflect the implications of such effects for future cash flows, leading to no future abnormal returns. Alternatively, if the market does not fully incorporate information about inflation gains, then stocks may be mispriced, leading to possible future abnormal returns. Accordingly, I predict that if investors overestimate (underestimate) the amount of unrecognized inflation gains in the current period, investors will be negatively (positively) surprised when these effects are realized in cash flows, leading to negative (positive) subsequent abnormal returns.

To test this prediction, I first conduct a portfolio-level return test. This test examines returns for portfolios constructed based on inflation-adjusted information, controlling for common risk

factors. More specifically, this test uses portfolios constructed based on inflation gains and investigates the intercepts (i.e., alphas or abnormal returns) obtained from time-series regressions of future returns in excess of the risk-free rate on the related-period Fama-French factor returns. This test reveals significant intercepts for the high and low inflation-based portfolios, with a zero-cost hedge strategy that results in a significant monthly abnormal return of 77–85 basis points, controlling for the Fama-French, momentum, and net operating assets factors. Thus, it appears that stock prices do not fully reflect the implications of inflation gains. I also find a negative association between inflation gains and future abnormal returns. This finding indicates that inflation gains are underestimated (overestimated) when current-period inflation gains are low (high), leading to the significantly positive (negative) abnormal future returns as the inflation gains turn into future cash flows. As a check, I also conduct a firm-level test that analyzes the extent to which an inflation-based trading strategy predicts firms' future annual abnormal returns. The firm-level test generates abnormal returns in a manner consistent with the portfolio-level test.

The findings show that, although investors take inflation gains into account, they do not do so correctly. In additional analysis, I conduct two tests to examine whether market pricing errors can be explained by how the stock market processes inflation information. I find that the documented ability to generate future abnormal returns and the negative association between inflation gains and future abnormal returns is consistent with an inflation adjustment argument. Specifically, I find that the presence of negative or positive future abnormal returns is consistent with investors not distinguishing monetary and nonmonetary assets when adjusting for inflation, leading to a loss of information because inflation affects these two classes of assets differently. This finding suggests that stock prices act as if investors "fixate" on aggregate assets when taking inflation into account, which leads to errors when adjusting for the effects of inflation.

Robustness tests show that there is no pattern in risk characteristics across inflation gains portfolios, and that an inflation-based factor is not a priced risk factor. These findings indicate that the abnormal returns I document are not attributable to an omitted inflation-based risk factor. Instead, these findings are consistent with abnormal returns stemming from inflation information being costly to obtain and process and, hence, are consistent with market efficiency under costly information (Beaver 1981).

To summarize, this study is the first to find that (1) unrecognized inflation gains are informative for predicting future cash flows, (2) inflation-based trading strategies are associated with significant abnormal returns, meaning that stock prices do not fully reflect the implications of inflation gains for future cash flows, and (3) stock prices act as if investors do not fully distinguish monetary and nonmonetary assets, which is necessary to determine the effects of inflation. Together, these results suggest that, contrary to prior work that concludes inflation-adjusted data are of little consequence to financial decisions, unrecognized inflation effects have significant economic implications, even during a period in which inflation is relatively low.⁴

Section II discusses related background and develops the study's main tests on cash flows and stock returns. Section III presents the research design. Section IV discusses data and sample statistics. Section V provides the main results. Section VI investigates how investors process inflation information. Section VII presents robustness tests. Section VIII concludes.

⁴ Although I find that inflation-adjusted amounts are informative for predicting future cash flows, I do not propose a shift to an inflation-adjusted reporting regime. Mandating an inflation-adjusted reporting regime in the U.S. may impose public- and firm-level costs that do not necessarily outweigh the benefits, especially when inflation is low. A normative evaluation of the trade-off associated with such a regime shift is beyond the scope of my study.

II. BACKGROUND AND MAIN TESTS

Inflation is a macroeconomic phenomenon that captures the decrease in purchasing power of a currency unit over time because of a general increase in the prices of goods and services (Beaver and Landsman 1983). The fundamental accounting assumption of a Monetary Unit assumes a stable currency (i.e., constant purchasing power) as the unit of record (SFAC 5, FASB 1984). Under U.S. GAAP, if the only activities of a firm are, for example, purchasing a parcel of land for \$100 fifty years ago and purchasing an additional parcel of land for \$100 one year ago, the firm recognizes land at \$200 in its financial statements. Assuming that the purchasing power changes over time, this results in a loss of information because the parcels were purchased with the same dollar amount, but at points in time with different purchasing power. Thus, although comparability is one of the qualitative characteristics of accounting information (SFAC 1, FASB 1978), mixing dollars from different periods distorts the Monetary Unit assumption and impairs comparability across firms and over time. This raises the question as to the economic effects of such inflation effects.

Implications of Inflation Information for Future Cash Flows

When a firm engages in a transaction, the opportunity cost of the transaction from the perspective of a representative investor is the fixed number of general consumption units the firm gives up. However, financial statements, which present nominal amounts, do not include information about the date of each transaction, and items are recognized in terms of dollars and not in terms of consumption units. Therefore, inflation creates a wedge between a recognized nominal amount and its cost in terms of consumption units. For a firm, nominal amounts mix dollars from periods with different dollar-consumption ratios, which impairs comparability over time. Across firms, nominal reporting further impairs comparability because there is large variation among firms' transaction dates and amounts when purchasing power is not constant.⁵

To extract information about transactions in terms of consumption units, financial statement items can be separated into two classes: monetary and nonmonetary. Monetary items are directly measured on the basis of a fixed number of dollars required for their settlement. Nonmonetary items represent either a historical cost or a right (obligation) to receive (deliver) services for which purchasing power is not constant. Under U.S. GAAP, financial statements do not take into account gains and losses on monetary items (e.g., debt) attributable to inflation. For example, although the erosion of monetary assets (liabilities) is a loss (gain) to the firm, such losses and gains are not recognized in U.S. GAAP-based financial statements. Financial statements also do not take into account the effects of inflation on nonmonetary items (e.g., PPE). For example, whereas inflation-adjusted amounts change over time because of inflation, the amounts recognized for gross PPE do not reflect such changes. The difference between earnings based on the two measures, $IGL = IAEarnings - NominalEarnings$, captures unrecognized inflation gains and losses. The variation in inflation-adjusted earnings, and hence in IGL , is a result of the level of and changes in inflation, and the difference across firms and over time in the structure of monetary and nonmonetary items. Thus, firms with similar nominal outcomes can differ in their inflation-adjusted out-

⁵ An example inspired from *The Economist's* "Big Mac index" (<http://www.economist.com/markets/bigmac/about.cfm>) illustrates how inflation can lead nominal accounting amounts to not reflect a transaction's opportunity cost in terms of consumption units. Suppose \$100 spent on a parcel of land 50 years ago could purchase 100 Big Macs 50 years ago, but because of inflation, \$100 can only purchase 40 Big Macs today. Because the \$100 spent 50 years ago is reported as \$100 in the current-period's nominal financial statements, the accounting system equates the original amount spent by the firm to the amount required to purchase 40 Big Macs today. That is, the accounting system effectively informs investors that the firm gave up an amount corresponding to 40 Big Macs to purchase the parcel of land. Yet the firm in fact purchased the land for the equivalent of 100 Big Macs. In inflation-adjusted terms, the \$100 spent 50 years ago is economically equivalent to \$250 ($=100 \times 100/40$) in today's dollars.

comes, and *IGL* can vary for a firm over time even when the nominal amounts remain constant.⁶

Inflation gains can turn into future cash flows in several ways, depending on firms' activities. Because the underlying economics of firms' activities are the same regardless of how these activities are reported in financial statements, if the nominal financial statements do not fully capture the effects of inflation today, such inflation effects are likely to be realized in future periods, thereby enabling the inflation effects to help predict future performance. The inflation effects can turn into future cash flows from operations (*CFO*) because higher unrecognized inflation gains accumulated in nonmonetary assets should result in higher future *CFO* when the assets are used (in the case of *PPE*) or sold (in the case of inventory). Further, because inflation is correlated with changes in specific prices, predicting higher future *CFO* from increases in the general price index is consistent with prior literature that shows increases in specific prices result in higher *CFO* (e.g., Aboody et al. 1999).⁷

There are many types of business activities for which inflation gains can turn into future cash flows. Consider, for example, the case of *PPE*. Suppose a parcel of land acquired for \$100 by the firm 50 years ago and, based on the purchasing power at that acquisition time, the firm gave up 100 Big Macs (i.e., one Big Mac—consumption unit—costs \$1). Suppose also that at the acquisition time, the firm expects to generate 10 percent inflation-adjusted annual yield by renting the land to another firm, i.e., it expects cash yield that is equivalent to ten Big Macs per year. Under a nominal reporting regime, firms' transactions are not linked to consumption units. Thus, when the purchasing power decreases over time, a particular number of consumption units in the past is equivalent to more cash in the present. Accordingly, in future periods, when the purchasing power of the dollar decreases because of inflation, the rent income is expected to increase, leading to higher *CFO*. Also, suppose a firm decides to use its land parcel for its own benefit rather than renting it to another firm. As the purchasing power of the monetary unit decreases over time, the use of the land parcel, which is nonmonetary, allows the firm to avoid paying an increasing amount to rent the parcel from another firm. This allows the firm to avoid cash payments that are increasing over time. Another example is inventory. Inventory is often sold several months or years after it is purchased, depending on its turnover, which is a function of the firm's operations and business cycle. The inventory accumulates inflationary gains that result in higher *CFO* upon the inventory sale. Thus, overall, I expect that inflation gains on *PPE* and inventory translate to higher future *CFO*.

If inflation gains, *IGL*, can turn into future *CFO*, then I predict a positive association between *IGL* and future *CFO*. I also expect inflation gains to turn into future cash flows over several years. First, inflation gains can be accumulated in nonmonetary assets (e.g., land, plant, and buildings) that usually have life cycles of several years. For example, unrealized inflation gains accumulated in a real estate property that is rented out can translate to *CFO* when the firm renews the lease or has another lessee beyond the one-year horizon. Second, inventory inflation gains can last longer than one year under LIFO, which many U.S. firms use, and when inventory turnover is longer than one year.

Note that inflation-adjusted accounting amounts are not fair value amounts. In contrast to fair value amounts, management discretion and subjectivity do not play a role when adjusting for

⁶ Inflation can lead to large differences between nominal and inflation-adjusted earnings even when inflation is low. For example, during a period of 6 percent average annual inflation in Israel, the median absolute difference between the two earnings amounts is 38 percent of nominal earnings. Also, during a period of 3 percent average annual inflation in the U.S., using the inflation-adjusted earnings for the U.S. sample, the equivalent ratio is 29 percent.

⁷ For example, untabulated results suggest that monthly inflation rates are highly and significantly correlated with monthly changes in major indices of nonmonetary assets (commodity and housing) over the past 59 years of available data in the Global Insight database, with Spearman and Pearson correlations of 0.7 on average.

inflation. When adjusting for inflation, I rely on the same measurement attribute underlying financial statements (mainly modified historical cost) and, thus, the adjustment procedure is objective. This is the same system of accounting as under the current nominal reporting regime, except that amounts are stated based on units of equal meaning. In particular, I adjust nominal amounts using a general price index to obtain a common monetary unit. The resulting (inflation-adjusted) amounts capture the implications of inflation for a representative investor who is interested in maintaining consumption units and is thus exposed to an average basket of goods and services.⁸

The Extent to which Stock Returns Reflect Inflation Information

Models in asset pricing often assume that an asset's price depends on the comovement of its payoff with marginal utility from consumption over time and across states of nature. Nominal reporting affects investors' assessment of firms' activities in terms of consumption units. If inflation gains, which capture firms' activities in terms of consumption units, are informative for predicting future cash flows, then this raises the question of whether the stock market fully incorporates such information when valuing firms' equity. Investors are likely to use information if it has implications for future cash flows. However, the use of inflation-adjusted data is more complicated than the use of nominal data (e.g., SFAC 5, FASB 1984). Furthermore, because classes of assets and liabilities are affected differently by inflation, and the inflation impact varies over time based on the structure and age of a firm's assets and liabilities, obtaining inflation-adjusted information (which is publicly unavailable) requires acquisition and processing costs that do not necessarily outweigh their benefits. These factors suggest that inflation accounting may reflect market efficiency under costly information (Grossman and Stiglitz 1980; Beaver 1981; Ball 1994).

If unrecognized inflation gains turn into future cash flows, then future returns should depend on how investors process information about these effects. Thus, I expect to find no subsequent abnormal returns if investors fully incorporate inflation information into their investment decisions. Alternatively, if investors do not fully incorporate information about inflation gains, stocks may be mispriced, leading to possible abnormal returns. Specifically, if investors overestimate (underestimate) inflation gains in the current period, then investors will be negatively (positively) surprised in the future when these gains turn into cash flows, leading to negative (positive) subsequent abnormal returns.

III. RESEARCH DESIGN

Inflation Adjustment Algorithm

To obtain accounting amounts that capture inflation effects, I develop an algorithm that converts nominal amounts into inflation-adjusted amounts on a firm-by-firm basis for a broad sample of firms.⁹ The algorithm uses a constant dollar approach to measure firms' activities as if purchasing power were constant. Further, it uses the same measurement attribute underlying nominal financial statements, restating the nominal amounts using an objective general price index.

⁸ Another reason that I adjust for inflation using a general price index, rather than specific indices for particular categories of assets is to be consistent with inflationary GAAP, which uses the general price index. Further, adjusting using specific assets requires not-readily available data regarding the composition of all assets within each asset class and the associated specific index that captures the price increase in the specific asset. Also, adjusting using specific industry indices is likely to introduce significant measurement error because a firm's composition of assets varies widely even for firms within the same industry. In fact, adjusting based on the general price index is likely to bias against finding significant abnormal returns and *CFO* predictability from using inflation gains.

⁹ Prior studies usually focus on particular countries or short periods because of data limitations. By using an algorithm to adjust nominal amounts for inflation I am able to analyze inflation effects on a broader sample.

The algorithm is described in the Appendix. In brief, the algorithm extracts inflation-adjusted information from GAAP financial statements in three main steps. First, it reconstructs, in inflation-adjusted terms, each firm's financial statements over the firm's life until the reporting date of interest. Because the distinction between monetary and nonmonetary items is key to distinguishing the effects of inflation on different accounting amounts, the algorithm separates monetary from nonmonetary assets and liabilities. Second, because monetary items are measured in financial statements based on the fixed number of dollars required for their settlement, the algorithm uses the monetary amounts from the financial statements as the inflation-adjusted amounts. In contrast, the algorithm adjusts nonmonetary items by estimating purchase dates and using information on asset life cycles and inventory turnover. Third, the algorithm estimates the wealth generated each period, in inflation-adjusted terms, by differencing two successive sets of assets and liabilities and using the clean surplus relation to incorporate other comprehensive income amounts, stock issues, and other events that can be affected by inflation over the period (e.g., dividends).¹⁰

Implications of Inflation Information for Future Cash Flows

If inflation-adjusted earnings are higher (lower) than nominal earnings in the current period, the unrecognized inflation gains (losses) should result in increased (decreased) future *CFO*. To test for the ability of inflation gains to predict future *CFO*, I extend the framework developed in Barth et al. (2001) to include inflation gains. Specifically, for horizons of one to four years, i.e., for $\tau = 1$ through 4, I estimate the following equation:

$$CFO_{t+\tau} = \theta + \beta \cdot IGL_t + \gamma_1 \cdot CFO_t + \gamma_2 \cdot \Delta AR_t + \gamma_3 \cdot \Delta INV_t + \gamma_4 \cdot \Delta AP_t + \gamma_5 \cdot DEPN_t + \gamma_6 \cdot OTHER_t + \mu_{t+\tau} \quad (1)$$

where IGL_t denotes inflation gains and losses for year t ; *CFO*, *INV*, *AR*, *INV*, *AP*, *DEPN*, and *OTHER* are cash flows from operations, accounts receivable, inventory, accounts payable, depreciation and amortization, and other accruals, respectively; and Δ denotes annual change.

In Equation (1), the coefficient on inflation gains, β , is the coefficient of interest. I predict a positive association between inflation gains and future *CFO* ($\beta > 0$). I also predict β to be positive for more than one year because assets can have life cycles longer than one year.¹¹ Following Barth et al. (2001), I predict the signs on γ_1 , γ_2 , γ_3 , γ_5 , and γ_6 to be positive and the sign on γ_4 to be negative.

By construction, the sum of *CFO* and disaggregated accruals equals nominal earnings, and thus the equation includes time t earnings. This allows me to test the incremental predictive effect of IGL_t beyond time t nominal earnings. I estimate Equation (1) using a pooled regression that includes observations cross-sectionally and over time, as well as using annual cross-sectional regressions. For the pooled regression, I base test statistics on regression residuals clustered by firm and year (Petersen 2009; Gow et al. 2010). I include industry fixed effects in the annual cross-sectional regressions, and I report the mean coefficients of time-series parameter estimates

¹⁰ Also, the adjustment procedure is consistent with inflationary GAAP that is currently active or was active in the past, e.g., International Financial Reporting Standards (IFRS): International Accounting Standards (IAS) 15 and 29, International Financial Reporting Interpretations Committee (IFRIC) 7; Israeli GAAP: Institute of Certified Public Accountants in Israel (ICPA) Statements 36 and 50; and U.S. GAAP: SFAS 33 and ASR 190.

¹¹ My motivation stems from whether there is a positive association between current-period inflation gains and future cash flows, thus I focus on the direction of the coefficient on inflation gains, rather than its magnitude. At the extreme there could be a one-to-one relation between inflation gains and future cash flows. Yet, predicting the magnitude of this coefficient is outside the scope of my study, and it depends on an array of factors and parameters that could vary widely depending on the assumptions regarding these parameters and for which data are unavailable (e.g., the life cycles of nonmonetary assets, the relation between inflation rate and the increase in specific asset prices and their benefits, and the composition of nonmonetary assets).

obtained from the cross-sectional regressions. I conduct tests on these cross-sectional estimates using the Fama and MacBeth (1973) t-statistic and two Z-statistics. I primarily rely on the robust double-clustered pooled specification because Fama-MacBeth t-statistic and the Z-statistics can be overstated if cross-sectional and time-series correlations exist in the data.¹²

The Extent to which Stock Returns Reflect Inflation Information

Portfolio-Level Approach

If mispricing exists, then a relation between inflation gains and subsequent returns is likely to be evident in portfolio-level returns. To test this prediction, the design focuses on the intercepts from portfolios constructed based on inflation gains. The estimated intercepts permit testing the ability of inflation information to explain systematic differences in the cross-section of stock returns, controlling for common risk factors (Fama and French 1993). In particular, I test whether the intercept for the low *IGL* portfolio is significantly different from the intercept for the high portfolio. To implement the test, I construct ten portfolios such that each period all firm-year observations with low (high) *IGL* are sorted into portfolio one (ten). I then calculate future monthly returns for each portfolio and estimate the following time-series equation at the portfolio level to obtain the portfolio intercepts:

$$R_{p,m} - R_{f,m} = \alpha_p + \beta_{p,MKTRF} \cdot MKTRF_m + \beta_{p,SMB} \cdot SMB_m + \beta_{p,HML} \cdot HML_m + \varepsilon_{p,m}, \quad (2)$$

where $R_{p,m}$ is the portfolio return for portfolio p in month m ; $R_{f,m}$ is the one-month Treasury bill rate; and $MKTRF_m$, SMB_m , and HML_m , are the Fama-French factor returns, where $MKTRF$ is the excess return on the market, and SMB and HML are constructed based on *MVE* and *BTM*, respectively. Because inflation gains are estimated annually, I align firms' *IGL* on monthly returns accumulated over the 12 months beginning three months after the fiscal year-end, to allow for dissemination of annual reports information and the associated annual inflation rate.

To test for hedge abnormal return from using an inflation-based trading strategy, I construct a zero-cost investment portfolio that longs the lowest portfolio (portfolio one) and shorts the highest portfolio (portfolio ten). I then regress this zero-cost portfolio's returns on the related-period factor returns. The intercept from this zero-cost hedge regression can be interpreted as a monthly abnormal return on a zero inflation-based hedge strategy that buys portfolio one and sells short portfolio ten.

I also sequentially add as controls the Carhart (1997) momentum factor (*UMD*) and a net operating assets factor (*FNOA*) following Hirshleifer et al. (2004). Hirshleifer et al. (2004) show that the ratio of net operating assets to lagged total assets, which they refer to as balance sheet bloat, is associated with future returns. I add this effect in all return tests to control for the possibility that a relation between Net Operating Assets (*NOA*) and inflation gains can affect the association between inflation gains and future returns. To do this, I first obtain *NOA* following Hirshleifer et al. (2004), and then form a *NOA*-based factor following the procedure described in

¹² Two points regarding this test. First, I focus on the inflation effects on future cash flows because cash flows receive considerable attention by the investment community and are of major importance in valuation of firms (e.g., Hackel et al. 2000; Barth et al. 2001). Cash flows are also more difficult to manipulate compared to earnings. However, inflation can also affect nominal earnings because inflation gains can translate into future nominal earnings in several ways (e.g., inventory holding gains). Second, with respect to the Z-statistics, both Z-statistics control for cross-sectional correlations but Z2 also partially corrects for potential upward bias in Z1 arising from lack of independence of parameters across the regression groups (Barth 1994). $Z1 = [(1/\sqrt{N}) \cdot \sum_{j=1}^N t_j] / \sqrt{[k_j / (k_j - 2)]}$, where t_j is the t-statistic for cross-sectional regression j , k_j is the degrees of freedom, and N is the number of cross-sectional groups; and $Z2 = \text{mean}(t) / [\text{stddev}(t) / \sqrt{(N-1)}]$, where $\text{mean}(t)$ and $\text{stddev}(t)$ respectively refer to the mean and standard deviation across the group estimates.

Fama and French (1993) in forming the *HML* and *SMB* factors. The *FNOA* factor is a factor-mimicking *NOA* portfolio.¹³

To the extent that inflation gains result in higher future cash flows, if investors correctly estimate inflation gains, then I expect the future abnormal return, α_p , to be insignificantly different from zero. Alternatively, if investors underestimate (overestimate) inflation gains, then I expect α_p to be significantly positive (negative), indicating positive (negative) subsequent abnormal returns as these inflation gains turn into higher- (lower-) than-expected future cash flows.

As depicted in Figure 1, I adjust for inflation year t amounts using the nominal amounts and inflation time-series rates until year-end t . Thus, inflation gains are known at year-end t and investors have this information before $t+1$ abnormal returns begin to accumulate. The association between current-period inflation gains and subsequent abnormal returns therefore depends on how the expected and unexpected components of inflation gains are estimated. Analogous to previous research, the source of the surprise that drives future returns, or the unexpected component of inflation gains at time t (and therefore the unexpected cash flows at $t+1$), is the difference between inflation gains that are estimated correctly versus inflation gains that are estimated without distinguishing monetary and nonmonetary assets.

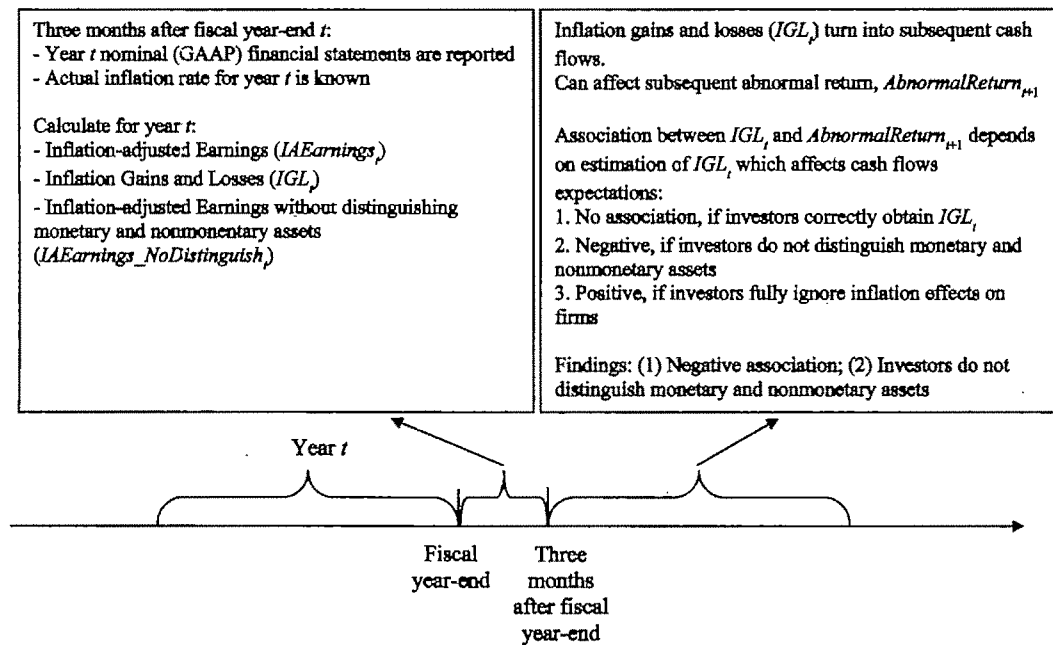
In other words, similar to the earnings surprise literature, the source of this study's surprise is the difference between the inflation adjustment that I estimate using the algorithm based on inflationary GAAP versus an adjustment that ignores the monetary-nonmonetary distinction.¹⁴ This type of analysis is analogous to that in Sloan (1996), who examines whether investors adequately distinguish the different persistence of the cash flows and accruals components of earnings in predicting future earnings.¹⁵ Also, my use of future abnormal returns to infer the

¹³ Following Hirshleifer et al. (2004), I obtain *NOA* as $NOA = RawNOA / TotalAssets_{t-1}$, where $RawNOA = Operating Assets - Operating Liabilities$; $Operating Assets = Total Assets (Compustat: AT) - Cash and Short-Term Investment (Compustat: CHE)$; and $Operating Liabilities = Total Assets (Compustat: AT) - Debt Included in Current Liabilities (Compustat: DLC) - Long-Term Debt (Compustat: DLTT) - Minority Interests (Compustat: MIB) - Preferred Stocks (Compustat: PSTK) - Common Equity (Compustat: CEQ)$. Next, I form a *NOA*-based factor following Fama and French (1993). At the end of each month, I sort all observations into two *NOA* groups, where group one (two) includes observations with low (high) *NOA*, and three book-to-market (*BTM*) groups, where group one (three) includes observations with low (high) *BTM*. I then construct six portfolios (*L/L*, *L/M*, *L/H*, *H/L*, *H/M*, *H/H*) from the intersections of the two *NOA* and three *BTM* groups, where the first letter in each of the *X/X* combinations refers to the *NOA* portfolio (Low, High) and the second letter refers to the *BTM* portfolio (Low, Medium, High). I then calculate monthly value-weighted returns on the six portfolios over the subsequent year, beginning three months after the fiscal year-end. *FNOA* is calculated each month as the average of the monthly returns on the three high *NOA* portfolios (*H/L*, *H/M*, and *H/H*) minus the average of the monthly returns on the three low *NOA* portfolios (*L/L*, *L/M*, and *L/H*).

¹⁴ Therefore, what leads to future abnormal returns is whether investors understand the differential effect of inflation on monetary versus nonmonetary assets, rather than investor understanding of expected versus unexpected inflation. This is different from investors' failure in the current period to distinguish expected and unexpected inflation. Specifically, when I estimate inflation-adjusted amounts, actual inflation is known because it is realized, as shown in Figure 1. Whether actual inflation is fully anticipated or fully unanticipated does not affect my predictions because investors use actual inflation, rather than its expected or unexpected component, to derive inflation gains. Further, the distinction between anticipated and unanticipated inflation is important when examining how current-period earnings changes explain contemporaneous stock price changes, under the notion that stock prices respond to the unanticipated inflation during the year—a setting that was widely used in the research design of inflationary accounting studies during the 1970s–1980s period. In contrast to this contemporaneous setting, my motivation and design are forward-looking, that is, in my study the events flow such that current-period (year t) inflation gains are estimated first, and only in the subsequent period ($t+1$) do these gains turn into cash flows. The subsequent returns thus do not arise from unanticipated inflation that affects year t inflation gains (see Figure 1).

¹⁵ To the extent that inflation gains are perfectly correlated over time, there may be no surprise component when these inflation effects turn into cash flows over time and thus there may be no theoretical link between inflation gains and future returns. However, when I calculate serial correlations in *IGL* over periods t and $t+1$, I find Pearson and Spearman correlations of 0.176 and 0.294, respectively. These correlations indicate that inflation gains show only weak persistence over time, as they have low serial correlation. Thus, there is no systematic relation among inflation gains over time. These results are consistent with my expectations, as inflation effects are likely to change over time because there is

FIGURE 1
Timeline



The figure plots the timeline for analyses throughout the study, and the link between current inflation gains and subsequent abnormal returns. Inflation gains and losses (IGL) are estimated after the fiscal year-end, using information from nominal financial statements and actual inflation rates as of the end of the year. IGL is defined as the difference between inflation-adjusted earnings and nominal earnings (scaled). Subsequent abnormal returns are accumulated over the year beginning three months after fiscal year-end. Because inflation gains affect the realization of subsequent cash flows, the accuracy of estimating inflation gains affects the accuracy of expectations regarding future cash flows. Subsequent returns are affected depending on actual versus expected inflation gains.

expected versus unexpected components is consistent with prior studies (e.g., Bernard and Thomas 1990; Sloan 1996). For example, similar to Sloan (1996), who infers the expected and unexpected persistence of accruals versus cash flows by examining future abnormal returns, I infer the expected and unexpected components of inflation gains (which lead to a future surprise when these gains affect future cash flows) by examining the patterns in future abnormal returns and inflation gains estimated with and without making the monetary versus nonmonetary distinction.

Firm-Level Approach

As a check on the portfolio-level analysis, I also conduct a firm-level test to discern whether it is possible to earn abnormal returns from inflation-adjusted information. In particular, each year, I form ten portfolios such that firms with the lowest (highest) IGL are sorted into portfolio one (ten). Then, for each portfolio-year, I calculate mean abnormal returns, accumulated over the

subsequent year. As before, I align firms' annual amounts to monthly returns in the next 12 months beginning three months after the fiscal year-end.

I examine abnormal returns using four metrics. These are returns adjusted for the value-weighted return on the market (market-adjusted), for the Fama-French factors (Fama-French adjusted), for the Fama-French and momentum factors (Fama-French-*UMD*), and for the Fama-French, momentum, and the net operating assets factors (Fama-French-*UMD-FNOA* adjusted). To obtain abnormal returns, I first calculate raw returns by annually compounding each firm's monthly returns. The market-adjusted return is calculated as the annually compounded raw return minus the annually compounded value-weighted return on all NYSE, AMEX, and NASDAQ stocks in CRSP. Next, I estimate the following time-series equation for each firm:

$$R_{i,m} - R_{f,m} = \kappa_i + \beta_{i,MKTRF} MKTRF_m + \beta_{i,SMB} SMB_m + \beta_{i,HML} HML_m + s_{i,m}. \quad (3)$$

I also estimate an equation similar to Equation (3) after sequentially adding as controls Carhart's (1997) momentum (*UMD*) and the above-defined balance sheet bloat (*FNOA*) factors. Estimation of Equation (3), and its augmented version with *UMD* and *FNOA*, over the sample period yields firm-specific betas, $\beta_{i,MKTRF}$, $\beta_{i,SMB}$, $\beta_{i,HML}$, $\beta_{i,UMD}$, and $\beta_{i,FNOA}$, which I winsorize at the top and bottom one percent. Finally, I obtain abnormal returns by subtracting from raw returns the product of a firm's betas and the respective factor returns, compounded annually.^{16,17}

IV. DATA AND SAMPLE STATISTICS

The sample covers U.S. firms with fiscal year-ends over the period 1984 to 2008, a period during which inflation was relatively low (average 3 percent). I obtain nominal accounting variables from the Compustat North America Fundamental Annual database, XPF Format. I obtain monthly raw stock returns from the CRSP Monthly Stock File, and I adjust these for delisting returns following Shumway and Warther (1999) and Beaver et al. (2007).¹⁸ I obtain the risk-free rate, the Fama-French and momentum factors, and portfolio returns from the Fama-French Portfolios and Factors dataset available through the Wharton Research Data Services (WRDS), and the consumer price indices data used in the inflation adjustment procedure from the Global Insight (DRI) dataset's Basic Economics Monthly Series (PRNEW).

To avoid extreme values obtained from deflating by a small denominator or from using a negative book value of equity, I delete observations with total assets, total revenues, or *MVE* lower than \$10 million, and observations with negative book value of equity. To align forecasting tests, and to verify that annual amounts are for a 12-month period, I delete firm-year observations if the firm's fiscal year changed during the year. To mitigate the effects of using penny stocks, I delete stocks with stock price lower than one dollar. The accounting variables are deflated by market

¹⁶ The firm-level test may lack power due to measurement errors in firm-specific factor betas and *IGL*. The portfolio test, in contrast, allows me to analyze the variation in the cross-section of expected returns. That is, rather than using firm-specific intercepts that depend on unknown firm-level characteristics, the portfolio test conditions on a pre-determined characteristic—inflation gains—and then identifies whether the inflation mispricing effects not explained by the factors vary with this characteristic. Thus, the portfolio test is less subject to this concern.

¹⁷ The return analyses involve the interaction of returns, which are nominal, with inflation-adjusted amounts. For comparability with inflation-adjusted amounts, the inflation effects are largely purged from the return metrics both by subtracting the risk-free interest rate from raw returns and by analyzing the incremental effect beyond that contained in the market factor. To check the extent to which the risk-free interest rate is associated with inflation, I compute correlations between the monthly time-series amounts of inflation and risk-free rates. The Spearman and Pearson correlations both are 0.42 and significant, consistent with the risk-free rate absorbing a large portion of the inflation effects from the return metrics. Further, I design the analyses to be on a cross-sectional basis, estimated either monthly or annually. This is because in any given period inflation-adjusted returns are perfectly correlated with nominal returns because inflation is the same for all firms (Beaver and Landsman 1983).

¹⁸ Inferences from all return tests are unchanged when I do not adjust for delisting returns.

value of equity (MVE) at the beginning of the year, MVE_{t-1} .¹⁹ To mitigate the effects of outliers, I winsorize all variables in Table 1 at the top or bottom one percentile of the deflated value in each year. To reduce measurement error in deriving inflation gains, either from nominal earnings or from using the algorithm, I delete observations in the top or bottom one percentile each year of deflated nominal earnings, inflation-adjusted earnings, and inflation gains. I use the same industry classifications as in Barth et al. (2010). The final sample comprises 64,597 U.S. firm-year observations.

The main explanatory variable of interest in the analysis is IGL , or $IAEarnings$ minus $NominalEarnings$. $NominalEarnings$ is Income before Extraordinary Items as reported in the financial statements, and $IAEarnings$ is nominal earnings restated on an inflation-adjusted basis using the algorithm described in the Appendix. To obtain data necessary to estimate Equation (1), I proceed as follows. For observations with fiscal years ending after July 15, 1988, I obtain CFO from the statement of cash flows (Barth et al. 1999). Specifically, CFO is net cash flow from operations less the accrual portion of extraordinary items and discontinued operations reported on the statement of cash flows, and total operating accruals ($ACCRUALS$) are calculated as $NominalEarnings$ minus CFO . The following are also calculated based on the statement of cash flows: change in accounts receivable (ΔAR), change in inventory (ΔINV), change in accounts payable and accrued liabilities (ΔAP), and depreciation and amortization ($DEPN$). If ΔAR , ΔINV , or ΔAP is missing, I calculate these items as I do for observations without statements of cash flows. The net of all other accruals, $OTHER$, is $OTHER = NominalEarnings + DEPN - (CFO + \Delta AR + \Delta INV - \Delta AP)$. For observations without a statement of cash flows available, that is, for observations with fiscal year ending before July 15, 1988, I derive CFO from accruals using information from the income statement and balance sheet (Dechow et al. 1995; Sloan 1996). To do so, ΔAR , ΔINV , and ΔAP are the change in the applicable balance sheet account (accounts receivable, inventory, and accounts payable plus accrued expenses, respectively), and $DEPN$ is depreciation and amortization as reported in the income statement. $ACCRUALS = \Delta AR + \Delta INV - \Delta AP - DEPN + OTHER$, where $OTHER = [(\Delta CA - \Delta CASH) - \Delta CL] - (\Delta AR + \Delta INV - \Delta AP)$; ΔCA , $\Delta CASH$, ΔCL are the respective changes in current assets, cash/cash equivalents, and current liabilities. Then, $CFO = NominalEarnings - ACCRUALS$. Additional variables I use include total assets ($TotalAssets$), net sales ($Revenues$), gross and net PPE ($GrossPPE$ and $NetPPE$), and MVE , calculated as fiscal year close price multiplied by common shares outstanding. The Appendix includes details on the Israeli data I use for the validation analysis.

Table 1 reports descriptive statistics for select variables. $TotalAssets$, $Revenues$, and MVE are presented in million U.S. dollars; $Book-to-Market$ and $TotalAssets / MVE_{t-1}$ are given as ratios; β_{LMKTRF} , β_{SMB} , β_{HML} , and β_{UMD} are the firms' Fama-French and momentum betas, estimated based on Equation (3); and all other variables are deflated by MVE_{t-1} . Panel A reports statistics based on pooled firm-year observations. It shows that the mean (median) MVE_{t-1} is \$2.17 billion (\$217 million). $TotalAssets$ are usually larger than MVE_{t-1} with a mean (median) ratio of 2.23 (1.30). The sign on the mean and median variables used in Equation (1) are generally consistent with prior research (e.g., Barth et al. 2001), with a positive sign on CFO , $NominalEarnings$, ΔAR , ΔINV , and $DEPN$, and a negative sign on ΔAP , $ACCRUALS$, and $OTHER$. Also, the means of $NonMonAssets$ (which consists of $NetPPE$, INV , and $Intangibles$) and $NetMonItems$ (monetary assets minus monetary liabilities), which are drivers in the inflation adjustment, respectively equal

¹⁹ I use MVE as the deflator in the cash flows prediction model for two reasons. First, it is consistent with prior research. Second, my goal is to control for scale differences using a measure that is uncorrelated with my algorithm to avoid spurious results driven from the error in the undeflated models being correlated with the regressors. In the derivation of inflation-adjusted earnings, the algorithm uses several accounting amounts of the firm, so deflating by accounting items (e.g., total assets) may introduce spurious correlations with the accounting variables when predicting cash flows.

TABLE 1
Descriptive Statistics

Panel A: Pooled Firm-Year Observations		Mean	Median	Std. Dev.	25th Pctl	75th Pctl
TotalAssets		3,874	287	34,430	89	1,153
Revenues		2,012	262	8,790	81	1,000
MVE _{t-1}		2,174	217	12,222	67	874
Book-to-Market		0.61	0.52	0.43	0.32	0.78
TotalAssets / MVE _{t-1}		2.23	1.30	4.04	0.71	2.41
CFO		0.12	0.09	0.19	0.04	0.18
ΔAR		0.03	0.01	0.13	0.00	0.04
ΔINV		0.02	0.00	0.09	0.00	0.02
ΔAP		-0.02	-0.01	0.13	-0.03	0.01
DEPN		0.07	0.04	0.09	0.02	0.09
OTHER		-0.08	-0.03	0.27	-0.09	0.01
ACCRUALS		-0.08	-0.04	0.20	-0.12	0.00
NominalEarnings		0.05	0.06	0.13	0.02	0.09
IAEarnings		0.03	0.04	0.17	0.00	0.08
IGL		-0.02	-0.01	0.09	-0.04	0.00
NOA		0.68	0.68	0.34	0.49	0.83
NetPPE		0.58	0.26	0.85	0.09	0.70
INV		0.26	0.10	0.46	0.01	0.30
Intangibles		0.18	0.04	0.36	0.00	0.18
OA		1.05	0.50	1.79	0.27	0.94
TotalLiabilities		1.33	0.60	2.13	0.23	1.45
ReExOCI		0.22	0.24	0.69	0.05	0.48
O		0.09	0.00	0.53	-0.03	0.00
CommonDividends		0.01	0.00	0.02	0.00	0.02
PreferredDividends		0.00	0.00	0.01	0.00	0.00
NewIssues		0.05	0.01	0.15	0.00	0.03
NonMonAssets		1.02	0.62	1.26	0.25	1.31
NetMonItems		-0.22	0.01	0.93	-0.44	0.24

(continued on next page)

Panel A: Pooled Firm-Year Observations

	Mean	Median	Std. Dev.	25th Pctl	75th Pctl
$\beta_{1,MKTRF}$	1.01	0.99	0.49	0.70	1.28
$\beta_{1,SMB}$	0.70	0.62	0.72	0.24	1.05
$\beta_{1,HML}$	0.19	0.33	0.81	-0.17	0.66
$\beta_{1,UMD}$	-0.16	-0.11	0.46	-0.35	0.08

Panel B: Means and Standard Deviations by Industry

	Total Assets		Net PPE		INV		Intangibles	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Chemicals	2,794	8,319	0.54	0.61	0.27	0.29	0.18	0.28
Computers	1,383	6,560	0.18	0.29	0.12	0.23	0.11	0.24
Durable Manufacturers	1,754	13,449	0.44	0.61	0.34	0.39	0.17	0.31
Extractive Industries	5,222	17,559	1.31	1.05	0.10	0.26	0.05	0.15
Financial Institutions	21,956	112,733	0.13	0.32	0.19	0.58	0.15	0.29
Food	3,699	9,700	0.69	0.86	0.36	0.52	0.26	0.46
Real Estate, Insurance	1,093	3,078	0.49	1.12	0.17	0.50	0.11	0.32
Mining, Construction	1,161	2,369	0.63	0.79	0.70	0.97	0.12	0.28
Other	19,742	90,020	0.54	0.69	0.28	0.46	0.19	0.39
Pharmaceuticals	2,441	8,445	0.16	0.23	0.08	0.15	0.10	0.23
Retail	1,446	5,411	0.59	0.75	0.57	0.68	0.17	0.36
Services	909	2,470	0.66	1.00	0.08	0.23	0.29	0.46
Textile/Print/Publish	1,410	3,380	0.63	0.70	0.35	0.43	0.24	0.45
Transportation	5,784	17,753	1.20	1.27	0.07	0.23	0.39	0.65
Utilities	4,645	7,831	1.81	1.01	0.08	0.12	0.08	0.27

(continued on next page)

	NonMonAssets		NetMonItems		PPE(t)		IAEarnings		IGL	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Chemicals	0.97	0.89	-0.22	0.67	61.95	16.31	0.03	0.16	-0.03	0.10
Computers	0.39	0.57	0.21	0.44	46.02	16.70	-0.01	0.16	-0.01	0.06
Durable Manufacturers	0.93	1.04	-0.11	0.72	60.49	16.50	0.03	0.16	-0.02	0.08
Extractive Industries	1.49	1.32	-0.63	0.95	59.95	13.59	0.02	0.22	-0.02	0.14
Financial Institutions	0.55	1.24	0.31	1.00	50.57	13.51	0.06	0.14	-0.01	0.07
Food	1.31	1.46	-0.47	0.97	64.51	15.42	0.05	0.15	-0.03	0.10
Real Estate, Insurance	0.83	1.66	0.15	1.26	48.17	14.48	0.02	0.21	-0.02	0.11
Mining, Construction	1.56	1.51	-0.59	1.06	54.95	16.13	0.03	0.20	-0.03	0.10
Other	0.99	1.11	-0.13	0.87	55.93	16.82	0.00	0.20	-0.02	0.09
Pharmaceuticals	0.32	0.45	0.12	0.33	56.68	16.96	-0.02	0.14	-0.01	0.05
Retail	1.33	1.40	-0.47	0.99	60.35	15.68	0.04	0.16	-0.02	0.09
Services	1.04	1.38	-0.25	1.01	51.32	15.84	0.03	0.18	-0.01	0.11
Textile/Print/Publish	1.21	1.22	-0.30	0.84	63.12	15.98	0.04	0.17	-0.03	0.10
Transportation	1.71	1.76	-0.79	1.28	58.73	16.09	0.01	0.24	-0.03	0.16
Utilities	1.99	1.16	-1.10	0.86	53.39	6.62	0.04	0.14	-0.04	0.12

The table presents summary statistics of select variables. *TotalAssets* (Compustat: AT), *Revenues* (Compustat: SALE), and *MVE* (market value of equity) are presented in million U.S. dollars; *Book-to-Market* (book value of equity divided by *MVE*), and *TotalAssets / MVE* are ratios; β_{LMKT} , β_{SMB} , β_{HML} , and β_{UMD} are firm-specific betas, obtained by regressing a firm's time-series monthly excess return on the Fama-French and momentum monthly factors over the sample period; *NOA* is net operating assets deflated by *TotalAssets*, and it is constructed following Hirshleifer et al. (2004); *PPE(t)* is stated in months; and, all other variables are deflated by MVE_{t-1} . *NominalEarnings* is Income before Extraordinary Items as reported in the financial statements (Compustat: IB); *IAEarnings* is nominal earnings restated to an inflation-adjusted basis using the algorithm; $IGL = IAEarnings - NominalEarnings$, and it captures unrecognized inflation gains and losses. The risk-free rate and the Fama-French and momentum factors are from the Fama-French Portfolios and Factors dataset available through WRDS. $NOA = RawNOA / TotalAssets_{t-1}$, where $RawNOA = Operating Assets - Operating Liabilities$, and it is obtained following Hirshleifer et al. (2004). Cash flows from operations (*CFO*) is extracted: directly from the statement of cash flows for firms with fiscal year ending after July 15, 1988, or it is derived for firms with fiscal year ending before July 15, 1988. Specifically, for observations with a statement of cash flows available, *CFO* is net cash flows from operating activities (Compustat: OANCF) less the accrual portion of extraordinary items and discontinued operations (Compustat: XIDOC); *ACCRUALS* is total operating accruals, calculated as $ACCRUALS = NominalEarnings - CFO$; ΔAR , ΔINV , ΔAP are the respective annual changes in accounts receivable (Compustat: RECCH), inventory (Compustat: INVCH), and accounts payable and accrued liabilities (Compustat: APALCH); *DEPN* is depreciation and amortization (Compustat: XDP); and *OTHER* is net of all other accruals, calculated as $NominalEarnings + DEPN - (CFO + \Delta AR + \Delta INV - \Delta AP)$. If Compustat items RECCH, INVCH, or APALCH is missing, ΔAR , ΔINV , or ΔAP is the change in the applicable balance sheet account, accounts receivable (Compustat: RECT), inventory (Compustat: INVT), or accounts payable (Compustat: AP) plus accrued expenses (Compustat: XACC), respectively. For observations without a statement of cash flows available, ΔAR , ΔINV , and ΔAP are the change in the applicable balance sheet account, accounts receivable (Compustat: RECT), inventory (Compustat: INVT), and accounts payable (Compustat: AP) plus accrued expenses (Compustat: XACC), respectively. For statements of cash flows available, $ACCRUALS = \Delta AR + \Delta INV - \Delta AP - DEPN + OTHER$, where *OTHER* represents accruals not directly captured by ΔAR , ΔINV , or ΔAP , i.e., *OTHER* equals $[(\Delta CA - \Delta CASH) - \Delta CL] - (\Delta AR + \Delta INV - \Delta AP)$, where ΔCA is the change in current assets (Compustat item: ACT), $\Delta CASH$ is change in cash/cash equivalents (Compustat: CHE), and ΔCL is change in current liabilities (Compustat: LCT). Then, $CFO = NominalEarnings - ACCRUALS$. *NonMonAssets* is the sum of *NetPPE*

(continued on next page)

(Compustat: PPENT), INV (Compustat: INVT), and *Intangibles* (Compustat: INTAN). Other variables used in developing the algorithm: $OA = TotalAssets - NonMonAssets$; *TotalLiabilities* (Compustat: LT + MIB); *ReExOCI* is retained earnings excluding other comprehensive income; O is other monetary items in stockholders' equity but not in retained earnings; *CommonDividends* (Compustat: DVC); *PreferredDividends* (Compustat: DVP); *NewIssues* is the change over the year in Common Stock (Compustat: CSTK), Preferred Stock (Compustat: PSTK), and Capital Surplus (Compustat: CAPS); and *NetMonItems* is monetary assets minus monetary liabilities. $PPE(t)$ is the expected remaining useful life in months for an average PPE item. Industries are based on Barth et al. (2010). The sample includes NYSE, AMEX, and NASDAQ U.S. stocks that are in the intersection of CRSP and Compustat, that have data available to calculate inflation gains and losses, and with fiscal year-ends between 1984 and 2008 ($n = 25$).

1.02 and -0.22 , with respective standard deviations of 1.26 and 0.93, indicating a high dispersion in these measures and in *IGL*. As expected given the broad sample, the mean (median) market beta of 1.01 (0.99). The mean (median) of *NominalEarnings* is 0.05 (0.06), which is higher than the mean (median) of *IAEarnings* of 0.03 (0.04). The mean (median) *IGL* is -0.02 (-0.01) with a standard deviation of 0.09.²⁰ The standard deviations of *IGL* and *IAEarnings* reveal a large variation in these measures.

Table 1, Panel B reports means and standard deviations for each industry by year. With respect to major nonmonetary items, the panel shows that the utilities, extractive, transportation, food, services, mining, and construction industries are the most capital intensive, with *NetPPE* varying between 0.63 and 1.81 in these industries. Panel B also shows that monetary liabilities are larger than monetary assets in most industries, and that *PPE(t)*, the expected remaining useful life (in months) for an average *NetPPE* item, is volatile, with the highest age of average assets in the food, textiles, printing, publishing, chemicals, durable, and retail industries (mean *PPE(t)* varies from 60.35 to 64.51 months in these industries). Even within industries, Panel B reveals high variation in monetary and nonmonetary items, which are drivers for inflation adjustments, suggesting high variation in inflation-adjusted earnings.

V. MAIN RESULTS

Implications of Inflation Information for Future Cash Flows

Table 2 reports results from the cash flows analysis. It shows that the coefficients on aggregated accruals are consistent with Barth et al. (2001). It also reveals that β is significantly positive based on the t-statistics and Z-statistics in the pooled and cross-sectional specifications for all four of the forecasting horizons considered.²¹ In particular, based on the pooled and cross-sectional specifications, β is significantly different from 0 and equals 0.05 and 0.07 (0.05 and 0.08; 0.05 and 0.11; 0.10 and 0.13), respectively, at the one- (two-, three-, four-) year horizon. This evidence suggests that inflation gains help predict future *CFO*.

The Extent to which Stock Returns Reflect Inflation Information

Table 3 reports results from the portfolio-level return analysis. The results show that five portfolio intercepts are significant, which indicates that forming portfolios based on inflation gains (*IG*) generates significant abnormal returns. The intercept on the lowest portfolio is 0.0049 ($p < 0.001$), whereas the intercept on the highest portfolio is -0.0036 ($p = 0.002$). The intercept of each portfolio can be interpreted as the monthly abnormal return from buying the specific portfolio and selling short a risk-free asset. The table also reveals that, controlling for the Fama-French factors, the monthly zero-cost hedge return for a portfolio constructed on the difference between the intercepts for portfolios one and ten equals 0.00850 and is significant ($p < 0.001$). The table also reveals that this significant abnormal hedge return holds when controlling for momentum (*UMD*) and net operating assets (*FNOA*) factors in addition to the three Fama-French factors, with significant monthly zero-cost hedge returns of 0.00826 ($p < 0.001$) and 0.00767 ($p < 0.001$), respectively. The findings indicate zero-cost abnormal hedge returns of 0.77 percent to 0.85 percent per month for using inflation information, controlling for common risk factors.

²⁰ As explained in the Appendix, *IGL* is normalized based on a reference point underlying the adjustment procedure. Because I adjust accounting amounts to be stated in terms of constant dollars to maintain the purchasing power at the estimated purchasing date of nonmonetary items, *IGL* is more frequently negative. Alternatively, *IGL* can be adjusted such that it is more frequently positive. The variation across firms and over time is unchanged under the two alternatives, and so are the cross-sectional results throughout the study.

²¹ p-values of 5 percent or less are considered statistically significant, and all significance levels are one-tailed where I have predictions and two-tailed otherwise.

TABLE 2
Implications of Inflation Information for Future Cash Flows

TABLE 2										
Implications of Inflation Information for Future Cash Flows										
	Prediction	θ	IGL	CFO	ΔAR	ΔINV	ΔAP	DEPN	OTHER	Adj.R ²
		?	+	+	+	+	-	+	?	
		CFO _{t+1}								
Pooled	Coeff.	0.02	0.05	0.52	0.25	0.14	-0.02	0.68	0.10	0.381
	t-stat	17.14	2.95	40.31	14.86	6.23	-1.03	31.06	7.73	
	p-value	<0.001	0.002	<0.001	<0.001	<0.001	0.152	<0.001	<0.001	
Cross-Sectional	Mean Coeff.	0.01	0.07	0.51	0.29	0.23	0.09	0.66	0.12	0.402
	FM t-stat	0.92	4.37	17.95	11.81	7.01	2.15	17.55	3.75	
	p-value	0.186	<0.001	<0.001	<0.001	<0.001	0.024	<0.001	<0.001	
	Z1	0.92	7.64	102.58	35.60	23.57	7.45	64.42	18.54	
	p-value	0.186	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	
	Z2	0.67	4.46	12.11	15.37	6.90	2.05	20.68	4.98	
	p-value	0.256	<0.001	<0.001	<0.001	<0.001	0.029	<0.001	<0.001	
		CFO _{t+2}								
Pooled	Coeff.	0.03	0.05	0.52	0.25	0.22	-0.06	0.74	0.10	0.307
	t-stat	20.51	2.77	34.00	12.53	8.46	-2.33	26.50	6.68	
	p-value	<0.001	0.003	<0.001	<0.001	<0.001	0.010	<0.001	<0.001	
Cross-Sectional	Mean Coeff.	0.00	0.08	0.51	0.27	0.26	-0.01	0.71	0.12	0.328
	FM t-stat	-0.15	3.06	27.96	10.48	7.74	-0.33	19.53	5.07	
	p-value	0.441	0.004	<0.001	<0.001	<0.001	0.373	<0.001	<0.001	
	Z1	-0.30	6.47	76.70	25.37	21.37	-0.73	53.92	15.52	
	p-value	0.384	<0.001	<0.001	<0.001	<0.001	0.238	<0.001	<0.001	
	Z2	-0.28	2.80	17.25	11.74	7.67	-0.30	21.57	6.73	
	p-value	0.392	0.006	<0.001	<0.001	<0.001	0.384	<0.001	<0.001	

(continued on next page)

TABLE 2 (continued)

Prediction	θ	IGL	CFO	ΔAR	ΔINV	ΔAP	DEPN	OTHER	Adj.R ²
	?	+	+	+	+	-	+	?	
<i>CFO_{t+3}</i>									
Pooled	Coeff.	0.04	0.52	0.29	0.29	-0.07	0.87	0.10	0.266
	t-stat	22.48	27.17	11.00	9.22	-2.41	25.57	5.08	
	p-value	<0.001	<0.001	<0.001	<0.001	0.008	<0.001	<0.001	
Cross-Sectional	Mean Coeff.	0.00	0.51	0.31	0.32	-0.03	0.82	0.12	0.295
	FM t-stat	-0.10	19.84	9.24	6.28	-0.98	19.97	4.89	
	p-value	0.461	<0.001	<0.001	<0.001	0.171	<0.001	<0.001	
	Z1	-0.25	58.30	23.14	20.93	-1.87	49.28	12.46	
	p-value	0.403	<0.001	<0.001	<0.001	0.040	<0.001	<0.001	
	Z2	-0.20	16.90	9.58	6.25	-1.21	19.43	6.16	
	p-value	0.422	<0.001	<0.001	<0.001	0.122	<0.001	<0.001	
<i>CFO_{t+4}</i>									
Pooled	Coeff.	0.06	0.56	0.34	0.25	-0.09	1.04	0.10	0.239
	t-stat	21.09	22.46	9.64	6.67	-2.15	23.49	3.71	
	p-value	<0.001	<0.001	<0.001	<0.001	0.016	<0.001	<0.001	
Cross-Sectional	Mean Coeff.	0.00	0.54	0.39	0.31	-0.04	0.93	0.13	0.270
	FM t-stat	0.04	15.63	9.35	6.87	-0.89	20.76	4.12	
	p-value	0.484	<0.001	<0.001	<0.001	0.193	<0.001	<0.001	
	Z1	-0.11	47.49	21.59	15.21	-1.91	43.71	10.49	
	p-value	0.457	<0.001	<0.001	<0.001	0.037	<0.001	<0.001	
	Z2	-0.12	14.74	9.33	6.71	-0.99	19.03	4.36	
	p-value	0.453	<0.001	<0.001	<0.001	0.168	<0.001	<0.001	

(continued on next page)

TABLE 2 (continued)

The table reports results from estimating pooled and cross-sectional regressions of the following equation:

$$CFO_{i,t+\tau} = \theta + \beta \cdot IGL_t + \gamma_1 \cdot CFO_t + \gamma_2 \cdot \Delta AR_t + \gamma_3 \cdot \Delta INV_t + \gamma_4 \cdot \Delta AP_t + \gamma_5 \cdot DEPN_t + \gamma_6 \cdot OTHER_t + \mu_{i,t+\tau}$$

for $\tau = (1, 2, 3, 4)$.

The pooled (cross-sectional) specification is estimated using double-clustering of regression errors by firm and year (annual cross-sectional regressions and including industry indicator variables that are omitted for brevity). CFO is cash flows from operating activities. $IGL = IAEarnings - NominalEarnings$, deflated by the lagged market value of equity, and it captures unrecognized inflation gains and losses. $NominalEarnings$ is Income before Extraordinary Items. $IAEarnings$ is nominal earnings restated to an inflation-adjusted basis using the algorithm. The control variables are defined in Table 1 and are based on Barth et al. (2001). Industries are based on the 15 industries in Barth et al. (2010). In the cross-sectional specification: Mean refers to the mean coefficient, averaged across the cross-sectional estimates; FM t-stat refers to the Fama and MacBeth (1973) t-statistic; Z1 = $[(1/\sqrt{N}) \cdot \sum_{j=1}^N t_j] / \sqrt{[k_j/(k_j-2)]}$, where t_j is the t-statistic for cross-sectional regression j ; k_j is the degrees of freedom, and N is the number of cross-sectional groups; and Z2 = $mean(t) / [stddev(t) / \sqrt{(N-1)}]$, where $mean(t)$ and $stddev(t)$ respectively refer to the mean and standard deviation across the group estimates. p-values for sign-predicted coefficients are obtained from one-tailed t-tests. The sample includes all NYSE, AMEX, and NASDAQ U.S. stocks that are in the intersection of CRSP and Compustat, that have data available to calculate inflation gains and losses, and with fiscal year-ends between 1984 and 2008 ($n = 25$).

TABLE 3
Abnormal Returns from Using Inflation Information: Portfolio Level

IGL Portfolio	Coefficient			t-statistic (p-value)			Adj. R ²	Mean IGL
	α_p	β_{MTRF}	β_{SMB}	β_{HML}	β_{MTRF}	β_{SMB}		
1	0.0049	1.033	0.776	0.242	40.14	21.62	0.893	-0.080
2	0.0046	0.986	0.772	0.226	44.97	25.24	0.915	-0.035
3	0.0042	0.992	0.723	0.233	45.56	23.82	0.914	-0.026
4	0.0041	1.008	0.745	0.191	50.08	26.53	0.929	-0.020
5	0.0030	1.004	0.671	0.145	46.84	22.43	0.917	-0.016
6	0.0017	0.982	0.621	0.217	39.99	18.11	0.884	-0.011
7	0.0001	0.982	0.507	0.278	39.48	14.62	0.871	-0.007
8	-0.0004	1.030	0.508	0.255	41.97	14.84	0.884	0.005
9	-0.0013	1.055	0.607	0.212	43.29	17.86	0.897	0.012
10	-0.0036	1.086	0.716	0.193	41.15	19.43	0.893	0.068
Fama-French Factors:								
Zero-cost hedge (1-10)	0.00850							
Add UMD, No FNOA:								
Zero-cost hedge (1-10)	0.00826							
Add UMD, Add FNOA:								
Zero-cost hedge (1-10)	0.00767							

(continued on next page)

TABLE 3 (continued)

The table reports results from estimating time-series monthly portfolio regressions. First, each firm-year observation accumulates return over the 12 months beginning three months after the fiscal year-end. Second, ten portfolios are constructed each period with the lowest (highest) *ICL* sorted into portfolio one (ten), rebalanced monthly. *ICL* is inflation-adjusted earnings minus nominal earnings (scaled), and it captures unrecognized inflation gains and losses. Third, average portfolio excess returns, $R_{p,m} - R_{f,m}$, are calculated each month *m* using all observations in that portfolio, and these monthly portfolio excess returns are regressed on the Fama-French factors (*MKT*, *SMB*, *HML*), sequentially adding as controls the Carhart (1997) momentum (*UMD*) and the net operating assets (*FNOA*) factors. The intercepts from these regressions are reported as α_p in the table. The magnitude and statistical tests conducted on the difference between the highest and the lowest portfolios are from regressing zero-cost investment hedge portfolio returns, which are obtained after longing the lowest portfolio and shorting the highest portfolio, on the related-period factors. The risk-free rate, $R_{f,m}$, is the one-month Treasury bill rate. Monthly raw stock returns are obtained from the CRSP Monthly Stock File, and are adjusted for delisting returns. The risk-free rate and the Fama-French and momentum factors from the Fama-French Portfolios and Factors dataset available through WRDS. *NOA* is net operating assets deflated by lagged total assets, constructed as described in Hirshleifer et al. (2004). *FNOA* is a factor-mimicking *NOA* portfolio, formed by first sorting each month *m* all observations into two *NOA* groups and three book-to-market (*BTM*) groups, and constructing six portfolios from the intersections of the two *NOA* and three *BTM* groups. I then calculate monthly value-weighted returns on the six portfolios over the subsequent year, beginning three months after the fiscal year-end. *FNOA* is then calculated each month as the average of the monthly returns on the three high *NOA* portfolios minus the average of the monthly returns on the three low *NOA* portfolios. *p*-values are obtained from two-tailed *t*-tests. The sample includes all NYSE, AMEX, and NASDAQ U.S. stocks that are in the intersection of CRSP and Compustat, that have data available to calculate inflation gains and losses, and with fiscal year-ends between 1984 and 2008 (*n* = 25).



Untabulated results from a trend regression of the intercepts on ordered portfolios indicate that the decline in intercepts across the ten portfolios is significant ($p < 0.001$), suggesting a significantly negative relation between inflation gains and subsequent abnormal returns.

Table 4 reports results from the firm-level return analysis. The table presents the annual mean abnormal returns, accumulated over the year subsequent to portfolio formation and aggregated over the lowest and highest *IGL* portfolios. Similar to the results from the portfolio-level analysis, the firm-level results show that inflation information generates significant abnormal returns, with statistically significant mean hedge abnormal returns over the sample period that vary between 8.8 percent and 9.8 percent per year across the four abnormal return metrics considered. The results also show that, consistent with the results in Table 3, the low portfolio consistently yields higher subsequent returns than the high portfolio, with positive abnormal returns to a hedge strategy in 22 to 23 of the 25 sample years.²²

Figure 2 plots the four abnormal return metrics accumulated over the year subsequent to portfolio formation, illustrating the dominance of the low portfolio (solid lines) over the high portfolio (dotted lines) on average in about 90 percent of the sample years. This reflects a time-consistent pattern in future abnormal returns.

Overall, the findings of significant future abnormal returns provide consistent evidence that inflation-adjusted information is not fully incorporated by the stock market.

VI. INVESTOR PROCESSING OF INFLATION INFORMATION

The association between current-period inflation gains (IGL_t) and subsequent abnormal returns ($AbnormalReturn_{t+1}$) depends on how investors estimate IGL_t . Figure 1 illustrates that IGL_t is estimated using information before $AbnormalReturn_{t+1}$ begin to accumulate. If investors correctly estimate IGL_t , then there is nothing unexpected with respect to the future realization of inflation gains in cash flows, such that no future abnormal returns are predicted (and hence no association between IGL_t and $AbnormalReturn_{t+1}$ is predicted). Alternatively, if investors completely ignore inflation, rendering as unexpected the entire future realization of inflation gains in cash flows, then it leads to a future positive surprise when inflation gains are realized. Such a surprise when ignoring inflation gains is commensurate with IGL_t because the higher the (ignored) inflation gains are, the more favorable are future cash flows relative to investors' expectations, leading to a predictable positive association between IGL_t and $AbnormalReturn_{t+1}$.

However, the previous section provides evidence of (1) the existence of future abnormal return, and (2) a negative association between IGL_t and $AbnormalReturn_{t+1}$. Thus, investors appear to neither correctly estimate nor completely ignore inflation gains. This raises the question as to whether, in attempting to adjust for inflation, investors make errors in doing so. Accordingly, I examine how inflation information is processed by investors. Prior studies indicate that investors "fixate" on aggregate amounts without distinguishing components of the aggregate amounts. For example, Sloan (1996) provides evidence consistent with investors "fixating" on aggregate earnings, failing to distinguish the different implications of the accrual and cash flow components of earnings for future performance. Similarly, because inflation affects monetary and nonmonetary assets differently, stocks prices will be affected if investors rely on aggregate amounts instead of distinguishing their different components. I therefore investigate how investors process the implications of inflation gains by examining whether "fixating" on aggregate amounts without distin-

²² The second column in Table 4 provides the annual inflation rate, calculated based on the PRNEW index from the Global Insight (DRI) dataset's Basic Economics Monthly Series. The inflation rate can be obtained from different data providers with slight differences, yet different measures are often highly correlated. For example, annual inflation rates over my sample period calculated from the Global Insight and World Bank datasets are very close to each other, with a high and significant correlation (correlation = 0.82).

TABLE 4
Future Abnormal Returns from Using Current-Period Inflation Information: Firm Level

Year of Portfolio Construction	Annual Inflation Rate	Market Adjusted Return			Add Factors			Add UMD			Add FNOA		
		Fama-French			Fama-French-UMD			Fama-French-UMD			Fama-French-UMD-FNOA		
		Low	High	Hedge	Low	High	Hedge	Low	High	Hedge	Low	High	Hedge
1984	0.0355	-0.12	-0.17	0.051	-0.09	-0.16	0.077	-0.06	-0.13	0.069	-0.06	-0.09	0.029
1985	0.0362	-0.02	-0.12	0.100	0.03	-0.05	0.080	0.04	-0.04	0.081	0.06	-0.04	0.105
1986	0.0074	0.03	-0.08	0.115	0.05	-0.05	0.104	0.04	-0.06	0.099	0.05	-0.05	0.104
1987	0.0447	0.03	-0.01	0.044	0.01	-0.03	0.041	0.02	-0.02	0.040	0.02	-0.02	0.046
1988	0.0428	-0.08	-0.12	0.040	0.01	-0.02	0.027	0.03	0.01	0.024	0.03	0.00	0.025
1989	0.0469	-0.02	-0.09	0.064	0.07	0.00	0.077	0.09	0.01	0.079	0.10	0.02	0.078
1990	0.0608	0.11	-0.01	0.127	0.04	-0.07	0.112	0.06	-0.06	0.117	0.07	-0.05	0.119
1991	0.0279	0.07	-0.09	0.157	0.00	-0.12	0.118	0.04	-0.09	0.130	0.04	-0.07	0.116
1992	0.0286	0.12	0.02	0.107	0.04	-0.04	0.085	0.06	-0.02	0.079	0.08	-0.01	0.090
1993	0.0257	0.05	-0.09	0.147	0.09	-0.05	0.143	0.09	-0.05	0.144	0.12	-0.03	0.155
1994	0.0271	0.07	-0.11	0.183	0.07	-0.11	0.186	0.09	-0.08	0.167	0.16	-0.06	0.214
1995	0.0251	-0.08	-0.14	0.058	-0.06	-0.12	0.061	-0.04	-0.09	0.051	-0.02	-0.07	0.047
1996	0.0330	-0.06	-0.08	0.020	-0.08	-0.11	0.032	-0.04	-0.06	0.022	-0.01	-0.05	0.040
1997	0.0147	-0.22	-0.37	0.151	-0.04	-0.13	0.093	0.01	-0.08	0.088	0.01	-0.08	0.088
1998	0.0158	0.36	-0.04	0.400	0.25	-0.15	0.402	0.34	-0.07	0.410	0.40	-0.03	0.425
1999	0.0279	0.26	0.16	0.099	0.21	0.00	0.204	0.23	0.02	0.203	0.28	0.04	0.234
2000	0.0338	0.42	0.22	0.204	0.29	0.12	0.175	0.31	0.13	0.176	0.34	0.13	0.208
2001	0.0129	0.08	0.00	0.077	0.07	0.01	0.052	0.09	0.03	0.063	0.13	0.03	0.106
2002	0.0248	0.56	0.30	0.260	0.30	0.09	0.209	0.29	0.09	0.204	0.31	0.10	0.206
2003	0.0186	-0.07	0.00	-0.074	-0.10	-0.01	-0.088	-0.06	0.01	-0.074	-0.06	0.02	-0.078
2004	0.0354	0.09	0.05	0.047	0.00	-0.04	0.043	0.03	-0.01	0.040	0.03	0.00	0.032
2005	0.0347	-0.04	-0.07	0.029	-0.01	-0.06	0.056	-0.01	-0.06	0.053	0.00	-0.06	0.055
2006	0.0242	-0.09	-0.06	-0.030	0.00	0.04	-0.044	0.04	0.07	-0.036	0.04	0.05	-0.013
2007	0.0446	0.04	-0.04	0.077	0.02	0.00	0.019	-0.01	-0.04	0.027	0.00	-0.04	0.039
2008	-0.0067	0.11	0.11	0.000	0.03	0.07	-0.043	-0.03	0.01	-0.047	-0.04	0.01	-0.048

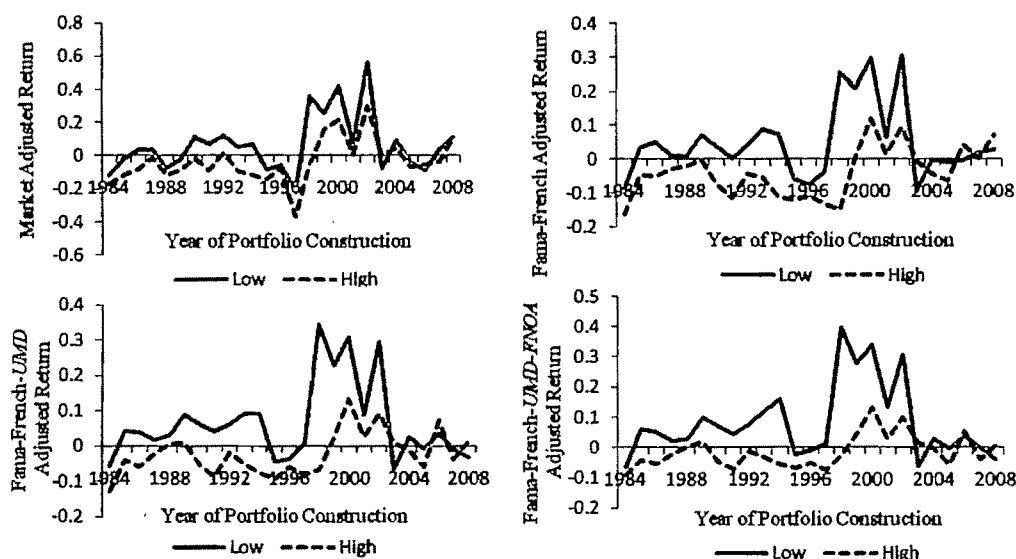
(continued on next page)

TABLE 4 (continued)

Year of Portfolio Construction	Annual Inflation Rate	Market Adjusted Return			Add Factors			Add UMD			Add FNOA		
		Low	High	Hedge	Fama-French Adjusted Return			Fama-French-UMD Adjusted Return			Fama-French-UMD-FNOA Adjusted Return		
					Low	High	Hedge	Low	High	Hedge	Low	High	Hedge
Mean				0.098			0.089			0.088			0.097
# Low > High				23.00			22.00			22.00			22.00
t-statistic				5.05			4.50			4.50			4.65
p-value				<0.001			<0.001			<0.001			<0.001

The table reports mean firm-level returns, calculated over the year subsequent to inflation-based portfolio formation. Specifically, observations with the lowest (highest) *IGL* are sorted annually into portfolio one (ten), where *IGL* is inflation-adjusted earnings minus nominal earnings (scaled), and it captures unrecognized inflation gains and losses. Then, for each portfolio-year, mean abnormal returns are accumulated over the subsequent year beginning three months after the fiscal year-end. Market Adjusted Return is the annually compounded raw return of the firm, $R_{i,t}$ minus the annually compounded value-weighted return on all NYSE, AMEX, and NASDAQ stocks in CRSP: Fama-French (Fama-French-UMD, Fama-French-UMD-FNOA) Adjusted Returns are $R_{i,t}$ minus the product of firm-specific betas and the respective Fama-French (Fama-French-UMD, Fama-French-UMD-FNOA) factors, where firm-specific betas are obtained by regressing the firm's time-series monthly excess return on the Fama-French (Fama-French-UMD, Fama-French-UMD-FNOA) factors. Fama-French factors are *MTRF*, *SMB*, and *HML*. *UMD* is the Carhart (1997) momentum factor. Raw stock returns are from the CRSP Monthly Stock File, adjusted for delisting returns. The risk-free rate and the monthly factors are from the Fama-French Portfolios and Factors dataset available through WRDS. *NOA* is net operating assets deflated by lagged total assets, constructed as described in Hirshleifer et al. (2004). *FNOA* is a *NOA*-based factor, formed as described in Table 3. The annual inflation rate is the annualized monthly consumer price index obtained from the Global Insight dataset's Basic Economics Monthly Series (PRNEW). *p*-values are obtained from two-tailed *t*-tests. The sample includes all NYSE, AMEX, and NASDAQ U.S. stocks that are in the intersection of CRSP and Compustat, that have data available to calculate inflation gains and losses, and with fiscal year-ends between 1984 and 2008 ($n = 25$).

FIGURE 2
Abnormal Returns from Using Inflation Information: Firm Level



The figure plots the annual mean firm-level returns reported in Table 4. The returns are accumulated over the year subsequent to portfolio formation, for the high and low decile portfolios formed each year based on *IGL*, such that observations with the lowest (highest) *IGL* are sorted into portfolio one (ten). *IGL* is inflation-adjusted earnings minus nominal earnings (scaled), and it captures unrecognized inflation gains and losses.

guishing the monetary and nonmonetary components can explain the abnormal returns findings. Distinguishing monetary and nonmonetary assets is critical in inflationary accounting because inflation affects the two classes of assets differently.

I conduct two tests to investigate this question. In the first test, I begin by using the algorithm to approximate inflation-adjusted earnings estimated without distinguishing monetary and nonmonetary assets (denoted as *IAEarnings_NoDistinguish*), by multiplying all assets by a constant as if there were no difference between the two classes of assets.²³ I then test for a systematic association between future returns and the difference between *IAEarnings_NoDistinguish* and *IAEarnings*—i.e., the inflation-adjusted earnings correctly estimated by distinguishing monetary and nonmonetary assets. Using the relationship between future returns and the expected surprise to infer how investors process information is similar to the approach used in prior studies (e.g., Bernard and Thomas 1990; Sloan 1996).

Table 5 reports the results. Panel A provides inflation-adjusted earnings with and without monetary and nonmonetary assets distinguished across *IGL* portfolios. It shows that not distinguishing monetary and nonmonetary assets leads to overestimation of inflation gains and losses (in

²³ Because of the clean surplus relation, this is equivalent to multiplying the income of all firms by the same constant, which is a function of the annual inflation rate, without taking into account differences in firms' asset composition.

TABLE 5
Future Returns and Investor Processing of Inflation-Adjusted Earnings

TABLE 5										
Future Returns and Investor Processing of Inflation-Adjusted Earnings										
IGL Portfolio	Panel A		Panel B				Panel C			
	Estimate of Current-Period Inflation-Adjusted Earnings, $IAEarnings_t$		Subsequent Abnormal Returns (from Table 3)				Regress α_p on $DIFF$			
	$IAEarnings$	$IAEarnings\_NoDistinguish$	$DIFF = IAEarnings - IAEarnings\_NoDistinguish$	α_p	t-stat	p-value	Slope	t-stat	p-value	R ²
1	-0.1705	-0.1935	0.0231 ^a	0.0049	4.37	<0.001	0.1507	4.74	0.001	0.738
2	-0.0205	-0.0249	0.0044	0.0046	4.78	<0.001				
3	0.0094	0.0077	0.0017	0.0042	4.46	<0.001				
4	0.0255	0.0303	-0.0049	0.0041	4.68	<0.001				
5	0.0310	0.0360	-0.0050	0.0030	1.68	0.093				
6	0.0351	0.0415	-0.0065	0.0017	1.62	0.106				
7	0.0385	0.0434	-0.0049	0.0001	0.11	0.913				
8	0.0484	0.0566	-0.0083	-0.0004	-0.38	0.703				
9	0.0744	0.0899	-0.0155	-0.0013	-1.26	0.207				
10	0.1876	0.2308	-0.0432 ^b	-0.0036	-3.10	0.002				

^a Positive news at $t+1$ when higher-than-expected inflation gains turn into cash flows.

^b Negative news at $t+1$ when lower-than-expected inflation gains turn into cash flows.

The table reports results from processing inflation information and its link to future returns.

Panel A reports current-period inflation-adjusted earnings, scaled by lagged market value of equity. $IAEarnings$ refers to inflation-adjusted earnings derived in a manner that distinguishes monetary and nonmonetary assets; this is the $IAEarnings$ measure used throughout the study and obtained using the algorithm. $IAEarnings\_NoDistinguish$ refers to inflation-adjusted earnings obtained using the algorithm without distinguishing monetary and nonmonetary assets.

Panel B reports the Fama-French Alphas, α_p from Table 3, which are $t+1$ abnormal returns obtained from portfolio regressions constructed on current-period inflation information, i.e., based on $IAEarnings$.

Panel C reports results from estimating a regression of α_p on $DIFF$ and an intercept (that is omitted), where $DIFF = IAEarnings - IAEarnings\_NoDistinguish$. IGL is inflation-adjusted earnings minus nominal earnings (scaled), and it captures unrecognized inflation gains and losses.

p-values are obtained from two-tailed t-tests. The sample includes all NYSE, AMEX, and NASDAQ U.S. stocks that are in the intersection of CRSP and Compustat, that have data available to calculate inflation gains and losses, and with fiscal year-ends between 1984 and 2008 ($n = 25$).

^a Positive news at $t+1$ when higher-than-expected inflation gains turn into cash flows.

^b Negative news at $t+1$ when lower-than-expected inflation gains turn into cash flows.

The table reports results from processing inflation information and its link to future returns.

Panel A reports current-period inflation-adjusted earnings, scaled by lagged market value of equity. $IAEarnings$ refers to inflation-adjusted earnings derived in a manner that distinguishes monetary and nonmonetary assets; this is the $IAEarnings$ measure used throughout the study and obtained using the algorithm. $IAEarnings\_NoDistinguish$ refers to inflation-adjusted earnings obtained using the algorithm without distinguishing monetary and nonmonetary assets.

Panel B reports the Fama-French Alphas, α_p , from Table 3, which are $t+1$ abnormal returns obtained from portfolio regressions constructed on current-period inflation information, i.e., based on $IAEarnings$.

Panel C reports results from estimating a regression of α_p on $DIFF$ and an intercept (that is omitted), where $DIFF = IAEarnings - IAEarnings\_NoDistinguish$. IGL is inflation-adjusted earnings minus nominal earnings (scaled), and it captures unrecognized inflation gains and losses. p-values are obtained from two-tailed t-tests. The sample includes all NYSE, AMEX, and NASDAQ U.S. stocks that are in the intersection of CRSP and Compustat, that have data available to calculate inflation gains and losses, and with fiscal year-ends between 1984 and 2008 ($n = 25$).

absolute value), such that *IAEarnings* are overestimated (underestimated) in the low (high) *IGL* portfolio, which results in an increase in the difference between *IAEarnings* and *IAEarnings_NoDistinguish* across *IGL* portfolios. Panel B provides the future abnormal returns (α_p) reported in Table 3 across *IGL* portfolios. Strikingly, the patterns in *IAEarnings* – *IAEarnings_NoDistinguish* (hereafter, *IAEarnings* – *IAEarnings_NoDistinguish* is referred to as *DIFF*) and α_p across the portfolios reveals that the two patterns are closely related, with the highest (lowest) future abnormal return arising when *IAEarnings* is at its highest (lowest) compared to *IAEarnings_NoDistinguish*. Specifically, not distinguishing monetary and nonmonetary assets leads to the greatest underestimation of inflation-adjusted earnings relative to their correctly estimated value (*DIFF* = 0.0231) in the lowest *IGL* portfolio, which is consistent with the return findings in that the lowest *IGL* portfolio has the highest and most significantly positive future abnormal return (α_p = 0.0049; $p < 0.001$). Thus, the results provide consistent evidence that the highest future abnormal return is obtained when *IGL* is most underestimated—i.e., when the future realization of inflation gains in cash flows is not fully expected, reflecting a future positive surprise when inflation gains turn into higher-than-expected cash flows. Investigating the highest *IGL* portfolio reveals similar consistency between inflation processing and future abnormal returns, with the lowest differential (*DIFF* = -0.0432) associated with the lowest and most significantly negative future abnormal return (α_p = -0.0036; $p = 0.002$). This consistency between the patterns in Panels A and B is also reflected in Panel C, which provides results from regressing α_p on *DIFF* and an intercept that is omitted from the table for brevity. The results show a significantly positive coefficient on *DIFF* (coefficient = 0.1507; $p = 0.001$).

In sum, the findings from this first test reveal that the error from not fully incorporating the effects of inflation is associated with the documented pattern in future abnormal returns. This test therefore provides evidence that the negative association documented between inflation gains and future abnormal returns is predictable and consistent with an inflation adjustment argument.

In the second test, I use a direct rational expectations framework similar to that employed in Mishkin (1983), Ball and Bartov (1996), and Sloan (1996). Specifically, to examine how inflation information is reflected in stock prices, I estimate the following nonlinear weighted least-squares system of equations:

$$\begin{aligned} CFO_{t+1} &= \delta_0 + \delta_1 \cdot IGL_t + \delta_2 \cdot IGL\_NoDistinguish_t + \delta_3 \cdot CFO_t + \delta_4 \cdot ACCRUALS_t + \zeta_{t+1} Return_{t+1} \\ &= \psi(CFO_{t+1} - \delta_0^* - \delta_1^* \cdot IGL_t - \delta_2^* \cdot IGL\_NoDistinguish_t - \delta_3^* \cdot CFO_t - \delta_4^* \cdot ACCRUALS_t) \\ &\quad + \mu_{t+1}, \end{aligned} \quad (4)$$

where the accounting variables are as defined above, and *Return*_{*t*+1} is the adjusted return accumulated over year *t*+1 and beginning three months after the fiscal year-end of year *t*. To control for heteroscedasticity, the first equation of the above system is scaled by the ratio of the residual variances of the two equations, where each equation's residual variance is obtained from an estimation of the given equation alone. Following my findings that *IGL*—constructed by distinguishing monetary and nonmonetary assets—predicts future *CFO*, I predict δ_1 (δ_2) to be significantly positive (insignificant). If stock prices correctly reflect the implications of inflation gains for future cash flows, then I expect δ_1^* (δ_2^*) to be significantly positive (insignificant), with the difference between the respective values of these two coefficients not significant. A significant difference between the coefficients on the inflation variables from the two equations would be evidence of investors not fully incorporating the implications of inflation information.

Table 6 reports the results. The table reveals that, as expected, δ_1 is significantly positive ($\delta_1 = 0.0815$; $p = 0.010$), δ_2 is insignificant ($\delta_2 = -0.0242$; $p = 0.289$), and the difference between δ_1 and δ_2 is significant ($p = 0.044$ for the test: $\delta_1 = \delta_2$), consistent with the ability of inflation gains and losses to predict future cash flows. However, the table also reveals that δ_1^* is insignifi-

TABLE 6
Stock Price Reaction to Information in Current Inflation Gains and Losses about Future Cash Flows

$$CFO_{t+1} = \delta_0 + \delta_1 \cdot IGL_t + \delta_2 \cdot IGL\_NoDistinguish_t + \delta_3 \cdot CFO_t + \delta_4 \cdot ACCRUALS_t + \zeta_{t+1}$$

$$Return_{t+1} = \psi(CFO_{t+1} - \delta_0^* - \delta_1^* \cdot IGL_t - \delta_2^* \cdot IGL\_NoDistinguish_t - \delta_3^* \cdot CFO_t - \delta_4^* \cdot ACCRUALS_t) + \mu_{t+1}$$

	Estimate	Asymptotic Standard Error	χ^2 -statistic	Asymptotic Pr > χ^2
δ_0	0.0587	0.0009	66.00	<0.001
δ_1	0.0815	0.0315	2.58	0.010
δ_2	-0.0242	0.0228	-1.06	0.289
δ_3	0.6912	0.0057	120.71	<0.001
δ_4	0.0825	0.0056	14.76	<0.001
Ψ	0.3814	0.0130	29.24	<0.001
δ_0^*	-0.2514	0.0131	-19.21	<0.001
δ_1^*	-0.4178	0.2723	-1.53	0.125
δ_2^*	1.6302	0.2044	7.98	<0.001
δ_3^*	0.7215	0.0493	14.62	<0.001
δ_4^*	0.6675	0.0521	12.81	<0.001
Tests				
$\delta_1 = \delta_2$			4.05	0.044
$\delta_1^* = \delta_2^*$			20.01	<0.001
$\delta_1 = \delta_1^*$			10.27	0.001
$\delta_2 = \delta_2^*$			64.71	<0.001

The table reports results from estimating a nonlinear weighted least squares system of equations to examine how inflation information is reflected in stock prices. CFO_{t+1} is cash flows from operating activities for year $t+1$, extracted directly from the statement of cash flows or derived from changes in balance sheet accounts as explained in Table 1. $ACCRUALS_t$ is total operating accruals, calculated as $ACCRUALS_t = NominalEarnings_t - CFO_t$, where $NominalEarnings$ is Income before Extraordinary Items as reported in the financial statements (Compustat: IB). IGL captures unrecognized inflation gains and losses and is calculated as $IGL = IAEarnings - NominalEarnings$, where $IAEarnings$ is nominal earnings restated to an inflation-adjusted basis using the algorithm in a manner that distinguishes monetary and nonmonetary assets; this is the inflation gains and losses measure used throughout the study. $IGL\_NoDistinguish$ approximates unrecognized inflation gains and losses without distinguishing monetary and nonmonetary assets and is calculated as $IGL\_NoDistinguish = IAEarnings\_NoDistinguish - NominalEarnings$, where $IAEarnings\_NoDistinguish$ refers to inflation-adjusted earnings obtained using the algorithm without distinguishing monetary and nonmonetary assets. All accounting variables are scaled. $Return_{t+1}$ is adjusted stock return for year $t+1$, accumulated over the year subsequent to year t and beginning three months after the fiscal year-end of year t . Raw stock returns are from the CRSP Monthly Stock File, adjusted for delisting returns. To control for heteroscedasticity, the first equation of the system is scaled by a scaling factor computed as the ratio of the residual variances of the two equations, where each equation's residual variance is obtained from estimating the equation alone. p-values are obtained from asymptotic Chi-squared tests. The sample includes all NYSE, AMEX, and NASDAQ U.S. stocks that are in the intersection of CRSP and Compustat, that have data available to calculate inflation gains and losses, and with fiscal year-ends between 1984 and 2008 ($n = 25$).

cant ($\delta_1^* = -0.4178$; $p = 0.125$), δ_2^* is significantly positive ($\delta_2^* = 1.6302$; $p < 0.001$), the difference between δ_1^* and δ_2^* is significant ($p < 0.001$ for the test: $\delta_1^* = \delta_2^*$), and the differences in the coefficients on the inflation variables across the equations are significant ($p = 0.001$ for the test: $\delta_1 = \delta_1^*$; $p < 0.001$ for the test: $\delta_2 = \delta_2^*$). Taken together, the findings from the second test are consistent with investors not fully incorporating the implications of *IGL* for future *CFO*, and with investors fixating on inflation gains and losses without distinguishing monetary and nonmonetary assets.

In sum, the two tests above use two different but complementary approaches to investigate how investors process inflation information in the cross-section of firms. The findings from these tests provide consistent evidence that the negative return pattern documented across portfolios of inflation gains is consistent with an inflation adjustment argument, whereby investors do not distinguish monetary and nonmonetary assets.²⁴

VII. ROBUSTNESS TESTS

Are the Abnormal Returns Attributable to Risk?

The return analyses provide evidence of mispricing. However, although the abnormal returns can be attributable to mispricing from inflation-adjusted information being costly to obtain and process, they may also be attributable to an omitted inflation-based risk factor. To verify the source of the abnormal returns, I first compare risk characteristics across *IGL* portfolios, and find that there is no pattern in risk characteristics across the portfolios.²⁵ I also conduct a two-step Fama-MacBeth test, which examines whether a risk factor is priced. This test uses an inflation-based factor and analyzes whether there is a positive risk premium to this factor in a cross-sectional regression analysis (Fama and MacBeth 1973; Fama and French 1992).

In the first step, I estimate time-series regressions at the portfolio level, based on the 25 portfolios constructed on *MVE* and *BTM* (Fama and French 1993):

$$R_{p,m} - R_{f,m} = \lambda_p + \beta_{p,MKTRF,m} \cdot MKTRF_m + \beta_{p,SMB,m} \cdot SMB_m + \beta_{p,HML,m} \cdot HML_m + \beta_{p,UMD,m} \cdot UMD_m + \beta_{p,FIGL,m} \cdot FIGL_m + \kappa_m, \quad (5)$$

where *FIGL* is a monthly factor-mimicking inflation information portfolio that I form based on *IGL* information in a similar way to how *FNOA* is constructed in Section V, following the procedure Fama and French (1993) use in forming *HML* and *SMB*. I focus on the 25 Fama-French portfolios because explaining their cross-sectional pattern in returns has attracted increasing interest in the literature (e.g., Lettau and Ludvigson 2001).²⁶ I use the time-series portfolio regressions to obtain the predicted factor loadings (betas) for each of the 25 portfolios, estimated using

²⁴ Note that there may be other channels that contribute to the negative association between *IGL* and future returns. For instance, investors may apply different rules to relate *IGL* to future cash flows and, thus, although future cash flows are higher when inflation gains are higher (Table 2), growth in realized cash flows could be lower than expected. In this case there would be a positive association between *IGL* and future cash flows but a negative relation between *IGL* and abnormal returns. Alternatively, investors may naively suppose that *IGL* turns into future cash flows in a one-to-one manner over some finite horizon, ignoring potentially important differences in the investment process associated with inflation gains. In this case inflation gains may generate future cash flows that are lower than what investors expect, which would lead to negative future returns. In this paper it is difficult to empirically investigate such channels without either a more structural model of the firms' investment opportunities or more finely disaggregated cash flow data that provide details about the source of each firm's cash flows from operating and investment activities; making several additional assumptions would make my inflation measure noisy at best. Importantly, the two tests I use to infer how investors process inflation information are similar to those used in several prior studies.

²⁵ For example, the median $\beta_{p,MKTRF}$ ($\beta_{p,SMB}$; $\beta_{p,HML}$; $\beta_{p,UMD}$) varies between 1.00 (0.59; 0.11; -0.096) and 1.05 (0.68; 0.46; -0.13), respectively, with no trend across *IGL* portfolios.

²⁶ More information about the 25 Fama-French portfolios (constructed on five *MVE* and five *BTM* portfolios) is at http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

five-year rolling windows that end at month m , with the requirement of at least ten portfolio-month observations in each window.

In the second step, I estimate cross-portfolio monthly regressions of portfolio excess returns for month $m+1$ on the predicted rolling betas, such that each regression pools predicted betas for the 25 portfolios, as follows:

$$R_{p,m+1} - R_{f,m+1} = \phi_{m+1} + \gamma_{MKTRF,m+1} \cdot \hat{\beta}_{p,MKTRF,m} + \gamma_{SMB,m+1} \cdot \hat{\beta}_{p,SMB,m} + \gamma_{HML,m+1} \cdot \hat{\beta}_{p,HML,m} \\ + \gamma_{UMD,m+1} \cdot \hat{\beta}_{p,UMD,m} + \gamma_{FIGL,m+1} \cdot \hat{\beta}_{p,FIGL,m} + \sigma_{m+1}, \quad (6)$$

where $\hat{\beta}_{p,MKTRF,m}$, $\hat{\beta}_{p,SMB,m}$, $\hat{\beta}_{p,HML,m}$, $\hat{\beta}_{p,UMD,m}$ and $\hat{\beta}_{p,FIGL,m}$ are the predicted portfolio betas estimated in the first step using data conditioned on month m . I aggregate and conduct tests on the estimates following the Fama and MacBeth (1973) procedure.

Table 7 reports the results. With respect to the Fama-French factors, the results are consistent with prior research, such as Petkova (2006) and Core et al. (2008). The mean estimated coefficients (p-values associated with the Fama-MacBeth t-statistics) on the market, size, book-to-market, and momentum factors equal -0.0058 , 0.0004 , 0.0038 , and 0.0001 (0.122 , 0.857 , and 0.033), respectively. Turning to the coefficient of interest, the mean estimated coefficient on $\hat{\beta}_{p,FIGL,m}$ across the monthly cross-sectional regressions, i.e., $\gamma_{FIGL,m+1}$, is -0.0015 with a p-value of 0.318 . The insignificance of the coefficient associated with the inflation gains factor, $\gamma_{FIGL,m+1}$, is inconsistent with an omitted risk factor associated with inflation. Rather, this result is consistent with abnormal returns being attributable to mispricing. Such mispricing does not necessarily imply market inefficiency, however, because inflation-adjusted information is potentially costly to obtain and process.

Possible Confounding Factors

In additional robustness tests, I check whether the abnormal return findings can be explained by factors that are known to explain returns in the cross-section. First, nominal leverage and nominal earnings-to-price ratio (E/P) are known to explain average returns. Although all of the return models in this study include as a control the size and book-to-market factors that are shown to absorb the roles of nominal leverage and E/P in explaining average returns (Fama and French 1992), I analyze the pattern in leverage and E/P across the *IGL* portfolios. I find that the respective median leverage and E/P are 0.6 and 0.04 for portfolio one and 0.61 and 0.05 for portfolio ten, with small variation and no systematic pattern in these ratios across portfolios. I also find small variation and no systematic pattern in past returns across the *IGL* portfolios, with a median past 12-month return of 0.10 and 0.12 for portfolios one and ten, respectively. Second, to examine whether industry concentration helps explain the results from the return tests, I examine the percentage of observations from each of the 15 industries in each of the ten inflation gains portfolios. The results suggest that no specific industry dominates, with industry concentration varying from 0.7 percent to 27.3 percent in the most extreme cases across the portfolios. Third, I examine the possible effect of industry-related omitted variables on the positive association between inflation gains and future cash flows by allowing not only the regression constant in Equation (1) to vary by industry, but also the other coefficients. To do so, I estimate Equation (1) by industry (including year fixed effects), and summarize the industry cross-sectional estimates using Fama and MacBeth (1973) t-statistics. The results reveal no change in inferences regarding the significantly positive association between inflation gains and future *CFO*, with cross-industry mean coefficients equal 0.05579 , 0.04341 , 0.0348 , and 0.07375 for the one- through four-year-ahead horizon, respectively.

Fourth, I further examine the robustness of the negative pattern in the return results in Tables

TABLE 7
Two-Step Fama-MacBeth Procedure: Results from Second Step
Monthly Cross-Portfolio Regressions

$$R_{p,m+1} - R_{f,m+1} = \phi_{m+1} + \gamma_{MKTRF,m+1} \cdot \hat{\beta}_{p,MKTRF,m} + \gamma_{SMB,m+1} \cdot \hat{\beta}_{p,SMB,m} + \gamma_{HML,m+1} \cdot \hat{\beta}_{p,HML,m} + \gamma_{UMD,m+1} \cdot \hat{\beta}_{p,UMD,m} + \gamma_{FIGL,m+1} \cdot \hat{\beta}_{p,FIGL,m} + \sigma_{m+1}$$

	Prediction	Mean	Std. Dev.	Standard Error	Fama-MacBeth t-statistic	p-value
Φ_{m+1}		0.0125	0.0561	0.0036	3.51	<0.001
$\gamma_{MKTRF,m+1}$		-0.0058	0.0592	0.0038	-1.55	0.122
$\gamma_{SMB,m+1}$		0.0004	0.0345	0.0022	0.18	0.857
$\gamma_{HML,m+1}$	+	0.0038	0.0321	0.0020	1.85	0.033
$\gamma_{UMD,m+1}$		0.0001	0.0611	0.0039	0.03	0.976
$\gamma_{FIGL,m+1}$	?	-0.0015	0.0243	0.0015	-1.00	0.318
Adj. R^2		0.510				

The table presents summary results from monthly cross-portfolio regressions in the second step of a two-step Fama-MacBeth (1973) procedure. In the first stage, portfolio time-series regressions are estimated using the 25 Fama-French portfolios (Fama and French 1993). Value-weighted portfolio excess return for month m for each portfolio p , $R_{p,m} - R_{f,m}$, is regressed on the contemporaneous monthly returns for the Fama-French (*MKTRF*, *SMB*, *HML*), Carhart (1997) momentum (*UMD*), and inflation-based (*FIGL*) factors, using five-year rolling windows that end at month m (requiring at least ten portfolio-month observations in each window), as follows:

$$R_{p,m} - R_{f,m} = \lambda_p + \beta_{p,MKTRF,m} \cdot MKTRF_m + \beta_{p,SMB,m} \cdot SMB_m + \beta_{p,HML,m} \cdot HML_m + \beta_{p,UMD,m} \cdot UMD_m + \beta_{p,FIGL,m} \cdot FIGL_m$$

+ κ_m . This results in five predicted rolling betas as of month m . In the second step, monthly cross-portfolio regressions of portfolio excess returns for month $m+1$, $R_{p,m+1} - R_{f,m+1}$, are regressed on the predicted rolling betas from the first step, such that each regression pools predicted betas for the 25 portfolios estimated in the first step using data conditioned on month m . The estimates are aggregated and tested following the Fama and MacBeth (1973) procedure. Fama-French and momentum factors are from the Fama-French Portfolios and Factors dataset available through WRDS. Fama-French 25 portfolio excess returns are from Ken French's website. $R_{f,m}$ is the one-month Treasury bill rate. The inflation-based factor, *FIGL*, is formed following Fama and French (1993), using the same procedure I use to form *FNOA*, as described in Table 3. p-values are obtained from one-tailed (two-tailed) t-tests when there is (no) prediction. The sample includes all NYSE, AMEX, and NASDAQ U.S. stocks that are in the intersection of CRSP and Compustat, that have data available to calculate inflation gains and losses, and with fiscal year-ends between 1984 and 2008 ($n = 25$).

3–5 by testing whether it is *IGL* and not the balance sheet bloat that causes the return distribution I find across *IGL* portfolios. Specifically, I sort observations into five *NOA* and five *IGL* portfolios (independent sorts) and obtain abnormal returns for the five-by-five, double-sorted portfolios. To assess the pattern across the portfolios, I also estimate a predicted trend regression model by regressing abnormal return on the *IGL* portfolio number (and an intercept). Untabulated results reveal that the negative return pattern across *IGL* portfolios persists in each of the *NOA* portfolios. The trend, which is the slope coefficient from the trend regression, is negative in all *NOA* portfolios (with a trend varies between -0.0056 and -0.0372 across the four abnormal return metrics I use), and significantly so in all cases except three, where the trend is marginally significant (with a p-value varies between 0.001 and 0.104 across the related four abnormal return metrics). These results provide additional evidence that the return distribution in Tables 3–5 is not caused by the balance sheet bloat. I also conduct an analysis wherein I repeat the return tests in Tables 3 and 4

after constructing *FNOA* as explained in Section III, but using ten *NOA* portfolios (instead of two portfolios as in Section III). Thus, *FNOA* is constructed based on the intersection of three *BTM* portfolios and the top and bottom deciles of *NOA*. This is because most of the negative return documented in the balance sheet bloat is concentrated in the top one to two deciles of *NOA*. Untabulated inferences from this analysis are unchanged relative to the findings in Tables 3 and 4. For example, the hedge return in Table 4 is positive in 21 of the 25 sample years, with a mean of 0.097 and $p < 0.001$.²⁷

VIII. CONCLUDING REMARKS

I hypothesize and find that inflation information in the current period affects future cash flows. Specifically, I find that unrecognized inflation gains turn into future cash flows from operations over the subsequent four years. I also find that investors do not fully incorporate such information into their investment decisions, in that I find significant abnormal returns for inflation-based trading strategies. I further find that the direction of the abnormal returns is consistent with investors not fully distinguishing monetary and nonmonetary assets. Robustness tests show that the documented abnormal returns are not attributable to differences in risk characteristics, and that an inflation-based factor is not a priced risk factor. These results are consistent with future abnormal returns being attributable to mispricing from costly information rather than to an omitted risk factor. Overall, the findings suggest that inflation has significant implications for performance and stock prices, even when inflation is relatively low.

My findings complement prior studies that primarily investigate stock pricing effects over a short event window or the contemporaneous year. A common theme from these studies is that inflation-adjusted data are inconsequential for making financial decisions. Watts and Zimmerman (1980) and Beaver et al. (1983) suggest that a potential interpretation of this conclusion is that stock prices are not affected by inflation-adjusted data because such information is new, and market participants have not yet learned how to analyze and process it (the “learning effect”). Also, Beaver and Landsman (1983) reason that the inflation-adjusted data are potentially too complex and unfamiliar to use. My study contributes to this literature by indicating that (1) investors do not appear to ignore inflation-adjusted data, which suggests that the learning effect has been partially realized over the past decades, and (2) investors do not fully incorporate inflation information, consistent with such information being complex and unfamiliar. Further, this study is the first to identify the mechanism that leads to the mispricing by linking stock returns to the underlying valuation fundamentals and revealing how inflation-adjusted information is processed by the stock market.

²⁷ I conduct two additional sensitivity tests. First, I estimate a model that pools the ten *IGL* portfolios and allows the intercepts for the extreme portfolios to vary using indicator variables. Second, I examine whether the firm-level return tests hold when using median abnormal returns. Untabulated results reveal unchanged inferences. Also, this study’s return findings are distinct from, and cannot be explained by, inflation illusion. The inflation illusion hypothesis (Modigliani and Cohn 1979) posits that highly levered firms are more undervalued because of investors’ failure to incorporate the gain accruing from purchasing power depreciation of nominal liabilities. Because the erosion of nominal liabilities leads to higher inflation gains, the direct effect from the inflation illusion hypothesis is higher (lower) future abnormal returns when inflation gains are high (low) as investors who suffer from inflation illusion are positively (negatively) surprised over future periods. However, despite this offsetting effect of the inflation illusion on the findings from the return analyses, my results are incremental to the inflation illusion effect in that I find that future abnormal returns are negatively related to inflation gains (see Section V). The inflation illusion hypothesis also posits that investors irrationally discount inflation-adjusted cash flows using nominal interest rates. Here, in contrast, I investigate how inflation directly affects cash flows, rather than how the cash flows are discounted.

APPENDIX

The appendix develops and validates an algorithm for incorporating inflationary effects into accounting amounts, using only publicly available information, by adjusting nominal to inflation-adjusted amounts on a firm-by-firm basis for a broad sample of firms.²⁸

Inflation Adjustment Algorithm

Financial statements can be restated using the balance sheet or the income statement.²⁹ I rely on the balance sheet to adjust the nominal financial statements.³⁰

A. Step 1: Adjustment of Nonmonetary Items

Nonmonetary items are linked to the dollar as of the year-end, but represent either a historical cost or a right (obligation) to receive (deliver) services for which purchasing power is not constant. I adjust these items as follows:

A.1. PPE: I use the PPE life cycle to adjust PPE. An asset's useful life is the period over which the entity expects to consume economic benefits from the asset. Assuming that accounting depreciation, on average, reflects an asset's useful life, the PPE life cycle is the average number of years from the asset's purchase until it is fully depreciated. I thus calculate the PPE life cycle as: $PPELifeCycle_t = (1/n) \cdot \sum_{i=t-\tau+1}^t [GrossPPE_i / PPE\ Depreciation]_i$, averaged over the four years prior to year-end t ($n = 4$). Next, I adjust Net PPE as follows: $adjNetPPE_t = NetPPE_t \cdot CPI_t / CPI_{t-\tau(t)}$, where adj refers to "adjusted;" t refers to the year t fiscal year-end; $\tau(t)$ is the period prior to fiscal year-end t , stated in annual terms and calculated as $\tau(t) = 0.5 \cdot PPELifeCycle_t$; and CPI denotes the Consumer Price Index.³¹ If $PPELifeCycle$ is negative, missing, or greater than the Compustat median limit of weighted expected useful life among different asset classes, which is calculated based on the expected maximum useful life of different PPE classes (e.g., U.S. Regulation 2003) varying between 20 years (e.g., Machinery and Equipment) and 50 years (e.g., Other Structures and Facilities), then I set it to the median life cycle calculated using the Compustat population over the sample period.³²

²⁸ An extended Appendix that includes further information, rationale, and illustrative details regarding the algorithm is available from the author upon request.

²⁹ The clean surplus relation makes the two approaches equivalent. This is because the income statement approach derives inflation-adjusted income before financing expenses by adjusting income statement amounts, whereas the balance sheet approach first calculates inflation-adjusted earnings using two successive balance sheets and then calculates inflation-adjusted financing expenses as the difference between net earnings and income before financing expenses. Inflation-adjusted financing expenses are the same if derived using the balance sheet or the income statement, resulting in same inflation-adjusted earnings under the two approaches.

³⁰ This is because (1) it avoids mistakes inherent in deriving $IAEarnings$ directly from the income statement, (2) it is more accurate because all transaction dates and income statement amounts are not necessary, and (3) because I focus on inflation-adjusted earnings, rather than inflation-adjusted revenues or gross profit, I can bypass reliance on further assumptions necessary to adjust the income statements (e.g., the timing of revenues over the year).

³¹ I multiply $PPELifeCycle$ by one-half because the life cycle is derived from gross, rather than net, PPE so the expected remaining useful life is one-half the gross PPE life cycle. Information about the exact transaction dates and amounts over the life of the firm is unavailable. Such information could help in estimating the exact purchasing date of each component of PPE and adjust it based on the associated vintage's purchasing power. Instead, I make a simplifying assumption that the PPE in place is acquired evenly over its life with the firm. That is, I adjust PPE using one-half of the Gross PPE life cycle such that the expected value of the remaining useful life is one-half of the life cycle obtained from Gross PPE. Also, note that because the adjustment is accurate to the monthly level, whereas t refers to annual amounts, τ is often a fraction (e.g., for an estimated purchase date of six months prior to fiscal year-end t , $\tau = 0.5$ and $NetPPE_t$ is adjusted using $CPI_t / CPI_{t-0.5}$).

³² Note that there can be alternative adjustment procedures depending on the assumptions used and the objectives underlying the adjustment. My objectives are to: (1) ensure consistency with actual inflationary GAAP; (2) obtain a sample of firms for which Compustat does not necessarily have available adjustment parameters (e.g., inventory and depreciation methods); and (3) develop a procedure that can be validated on firms in another country. For example, Davidson et al. (1976) requires data on the depreciation method. Requiring data about the inventory and depreciation methods would reduce my sample considerably, because such U.S. data are unavailable for about 40 percent of the observations.

A.2. Inventory: I use the inventory turnover, IT , ratio to adjust inventory, using the ratio $COGS/Inventory$, where $COGS$ is the Cost of Goods Sold. Year-end t inventory turnover is calculated as: $IT_t = COGS_t / [(INV_t + INV_{t-1}) / 2]$. If, e.g., $IT_t = 2$, the firm invests in inventory twice a year so the average inventory is six months old. In expectation, year t inventory will have a remaining life of $12 / (2 \cdot IT_t)$ when stated in months, or $\kappa(t) = 1 / (2 \cdot IT_t)$ when stated in years. Thus, I adjust inventory as follows: $adjINV_t = INV_t \cdot CPI_t / CPI_{t-\kappa(t)}$. If $COGS$ or INV are missing or negative, IT is set to the median IT of the Compustat population over the sample period.

A.3. Intangibles: I calculate the intangibles' remaining life for time t , denoted as $\omega(t)$, as the ratio of intangibles to the amortization of the related intangibles at time t . I assume that, in expectation, the number of years prior to the transaction generating the intangibles equals the remaining years until the amount of intangibles is fully reserved, and thus I adjust intangibles using the price index as of the expected value of the original transaction date, or $adjIntangibles_t = Intangibles_t \cdot CPI_t / CPI_{t-\omega(t)}$. I set intangibles' remaining life to the median remaining life of intangibles for the Compustat population over the sample period if it is negative, missing, or greater than firms' common weighted useful life of different intangibles classes, which is calculated based on the useful life of different intangibles classes varying between 2 and 40 years (e.g., patents) and between 20 and 40 years (e.g., goodwill). Also, according to SFAS 142 (effective in 2002), goodwill and other intangible assets no longer have a defined life for amortization but instead are tested annually for impairment. Because the algorithm uses amortization based on the pre-SFAS 141/142 period, it uses parameters obtained from the Compustat population to adjust the years that follow. I repeat all analyses without amortizing the years subsequent to 2002, and the inferences are unchanged.

A.4. Common Stock, Preferred Stock, and Capital Surplus: These items, which are included in shareholders' equity and represent purchasing power as of the stock issue dates, consist of two layers: (1) all stock issues from a firm's establishment through $t-1$, and (2) new equity issues occurring in year t (this layer can include several sub-layers, one from every equity issue that occurred over the year). I assume that equity issues are distributed uniformly over the year. To state amounts in constant dollars as of the reporting date, I begin by adjusting the first layer to derive retained earnings for both year $t-1$ and year t . In constant dollars as of t year-end, the adjusted amount of the first layer in $t-1$ is equal to the amount in t for calculating year t adjusted earnings. Using this two-layer process allows one to adjust earnings without having information about all the preferred and common stock issue dates and amounts from firms' incorporation dates until $t-1$. Thus, the following amount, which corresponds to the first layer and provides $t-1$ equity, appears in any two consecutive retained earnings and is used to extract inflation-adjusted earnings: $adjE_{t-1} = [CommonStock + PreferredStock + CapitalSurplus]_{t-1} \cdot CPI_t / CPI_{t-1}$. For the second layer, I obtain adjusted new issues during the year, $AdjNewIssues_t$, by calculating new issues, $NewIssues_t = [CommonStock + PreferredStock + CapitalSurplus]_t - [CommonStock + PreferredStock + CapitalSurplus]_{t-1}$, and adjusting this amount using one-half year's change in CPI , under the assumption that new issues occur uniformly throughout the year.

A.5. Other Monetary Items in Stockholders' Equity but not in Retained Earnings (O): Because earnings are obtained from the difference in retained earnings between two successive periods (adjusted for dividends and capital changes), it is necessary to exclude items that violate the clean surplus relation (e.g., Employee Benefit Trust) from inflation-adjusted retained earnings. This component is assumed to be monetary and is calculated as $O_t = TotalAssets_t - TotalLiabilities_t - ReExOCI_t - CommonStock_t - PreferredStock_t - CapitalSurplus_t$, where $ReExOCI_t$ is per A.6 below.

A.6. Retained Earnings Excluding Other Comprehensive Income ($ReExOCI$): It is critical to maintain the clean surplus relation when deriving earnings. Accordingly, I obtain nominal and

inflation-adjusted Retained Earnings Excluding Other Comprehensive Income. The inflation-adjusted amount is required because *IAEarnings* is derived using the two-period difference in inflation-adjusted *ReExOCI*. The nominal amount is used to derive *O* (per A.5.) as follows: $ReExOCI = \text{Retained Earnings (Compustat: RE)} - \text{Accumulated Other Comprehensive Income (Compustat: ACOMINC)}$. The inflation-adjusted *ReExOCI* as of year t , $adjReExOCI_t$, is derived by using the relation that total assets equal total liabilities plus shareholders' equity, and by stating all balance sheets amounts in constant dollars, where monetary (nonmonetary) items are not (are) adjusted: $adjReExOCI_t = adjINV_t + adjNetPPE_t + adjIntangibles_t + OA_t - adjE_{t-1} - adjNewIssues_t - O_t - TotalLiabilities_t$ (where, as above, $adjE_{t-1} = [CommonStock + PreferredStock + Capital Surplus]_{t-1} \cdot CPI_t / CPI_{t-1}$). Total liabilities are treated as monetary. I treat as monetary other assets (*OA*) that are not directly adjusted, and derive them as a residual value, using the relation that total assets equal total liabilities plus shareholders' equity, as follows: $OA_t = TotalAssets_t - INV_t - NetPPE_t - Intangibles_t$.

A.7. Other Comprehensive Income and Other Items Affecting Retained Earnings without Directly Affecting Net Income (*OtherInReExOCI*): This item is used in the equation that derives *IAEarnings*. Two types of exclusions are subtle, yet necessary for the accounting identities to hold and thus for the accuracy of the algorithm. First, because *IAEarnings* is obtained using the two-period difference in *adjReExOCI*, dividends must be included in the adjustment. Second, all transactions that are neither part of Other Comprehensive Income nor part of Net Income need to be excluded (e.g., Net Issues of Common Stock under Employee Plans; Purchases and Sales of Treasury Stocks under Employee Plans). Because these exclusions are the result of transactions occurring at the year-end, I treat them as monetary. These amounts are calculated as: $OtherInReExOCI_t = ReExOCI_t - ReExOCI_{t-1} - NetIncome_t + CommonDividends_t + PreferredDividends_t$.

A.8. Dividends: Because dividends are usually paid quarterly, the adjusted common and preferred dividends, *adjCommonDividends* and *adjPreferredDividends*, are adjusted assuming these payments are distributed uniformly over the year.

B. Step 2: Treatment of Monetary Items

Monetary assets and liabilities are measured on the basis of a fixed number of dollars required for their settlement. Thus, nominal monetary amounts are already stated in terms of constant purchasing power and, accordingly, I treat monetary items as equal to their recognized nominal amounts. The following are considered monetary: Cash, Short-Term Investments, Total Receivables, Total Liabilities, and assets not directly treated as nonmonetary assets (*OA*). The inclusion of *OA* implicitly treats unconsolidated but wholly-owned subsidiaries as monetary, consistent with Bernard and Hayn (1986).

C. Final Step: Derivation of Inflation-Adjusted Earnings

Inflation-adjusted earnings, *IAEarnings*, are calculated as follows:

$$IAEarnings_t = [adjReExOCI_t - adjReExOCI_{t-1}] + adjCommonDividends_t + adjPreferredDividends_t - OtherInReExOCI_t - adjExtraordinaryItems_t.$$

I obtain $adjReExOCI_{t-1}$ analogously to $adjReExOCI_t$ (see A.6. above), except that in this case (1) I adjust the accounting amounts reported for year $t-1$ to the purchasing power as of t year-end, and (2) I do not subtract $adjNewIssues_{t-1}$ because it is already part of $adjE_{t-1}$ as the new issues

during $t-1$ are part of the $t-1$ equity amount.³³ To reduce measurement error from the adjustment procedure, I delete observations each year in the top and bottom percentiles of the MVE_{t-1} deflated difference between *IAEarnings* and *NominalEarnings*. Because I investigate the behavior of *IAEarnings* versus *NominalEarnings* and because *NominalEarnings* refers to Net Income Excluding Extraordinary Items, I exclude extraordinary items when deriving *IAEarnings* to make the two earnings measures comparable. I assume that extraordinary items, if any occur, are distributed uniformly over the year and thus are adjusted using one-half year's change in the price index; these items are denoted as *adjExtraordinaryItems*.³⁴

External Validation of the Algorithm

To provide evidence on the external validity of the algorithm, I test the algorithm on a sample of Israeli firms. Until 2003 Israeli firms were required to recognize financial statements in inflation-adjusted terms and disclose in footnotes the same financial statements in nominal terms, and similar to the U.S., the inflation rate in Israel over the past decade was relatively low. In the validation analysis, I examine the extent to which nominal earnings derived by the algorithm, *NominalEarnings*^{Model}, approximates disclosed nominal earnings, *NominalEarnings*^{Actual}, by estimating the equation: $NominalEarnings^{Model} = \alpha + \beta \cdot NominalEarnings^{Actual} + \varepsilon$. If the algorithm does a good job translating earnings from one measurement basis into the other, I predict the intercept to be equal to 0 and the slope to be equal to 1. Thus, I conduct the tests: $H_0^{\alpha}: \alpha = 0$ against $H_1^{\alpha}: \alpha \neq 0$, and $H_0^{\beta}: \beta = 1$ against $H_1^{\beta}: \beta \neq 1$. To do so, I hand-collect data from Israeli firms' annual nominal and inflation-adjusted financial statements over the 1995–2003 period for 81 randomly selected firms listed on the Tel-Aviv 100 index. This index comprises the 100 firms with the highest *MVE* and accounts for more than 80 percent of the total market's capitalization. The 81 firms that I sample account for 86.63 percent of this index's total market capitalization as of December 21, 2005.

After implementing the algorithm and requiring the same restrictions as with the U.S. data, the inflation-adjusted Israeli sample includes 503 firm-year observations. Also, because footnotes are not always attached to the financial statements, causing nominal footnote disclosures to not always be available, I randomly select 50 firms and gather nominal information, when such footnotes are available. Monthly CPI and exchange rate data are obtained from the Israeli Central Bureau of Statistics. The Israeli sample reflects a median firm size of \$220 million. The mean and median values of the difference between actual (i.e., reported) inflation-adjusted earnings and

³³ Specifically, $adjReExOCI_{t-1} = adjINV_{t-1} + adjNetPPE_{t-1} + adjIntangibles_{t-1} + adjOA_{t-1} - adjE_{t-1} - adjO_{t-1} - adjTotalLiabilities_{t-1}$, where: $adjINV_{t-1} = INV_{t-1} \cdot CPI_t / CPI_{t-1-\kappa(t-1)}$; $adjNetPPE_{t-1} = NetPPE_{t-1} \cdot CPI_t / CPI_{t-1-\tau(t-1)}$; $adjIntangibles_{t-1} = Intangibles_{t-1} \cdot CPI_t / CPI_{t-1-\omega(t-1)}$; $adjOA_{t-1} = OA_{t-1} \cdot CPI_t / CPI_{t-1}$; $adjO_{t-1} = O_{t-1} \cdot CPI_t / CPI_{t-1}$; as above, $adjE_{t-1} = [CommonStock + PreferredStock + CapitalSurplus]_{t-1} \cdot CPI_t / CPI_{t-1}$; $adjTotalLiabilities_{t-1} = TotalLiabilities_{t-1} \cdot CPI_t / CPI_{t-1}$; and $\kappa(t-1)$, $\tau(t-1)$, and $\omega(t-1)$ refer to the period (stated in years) from which the lagged nonmonetary assets *INV*, *NetPPE*, and *Intangibles*, respectively, are adjusted.

³⁴ With respect to the derivation of *IGL*, there is a normalization based on a reference point underlying the adjustment procedure. Specifically, accounting amounts can be adjusted to be stated based on either constant dollars to maintain transactions in purchasing power, or current dollars to maintain transactions in consumption units. In the cross-section, the variation in *IGL*, rather than its level, is informative for explaining variation across firms, and the two approaches are equivalent when intercepts are added to the tests. I choose to adjust for constant dollars, leading *IGL* to be more frequently negative. Alternatively, *IGL* can be adjusted such that it is more frequently positive but the variation across firms and over time is unchanged. Accordingly, if the prediction model is $CF_{t+1} = a + b \cdot IGL_t + X_t + \eta_{t+1}$, where X is a vector of additional explanatory variables (conditioned on the time t information set), analyses throughout the study pertain to the parameter b , which is invariant to the reference point underlying the measurement system. The intercept, a , varies with the measurement system but is not a parameter of interest in my prediction analyses. Accordingly, the research design throughout my study includes intercepts in all cross-sectional tests and focuses on the coefficient on *IGL*.

nominal earnings, IGL^{Actual} , are -0.02 and -0.01 , with a standard deviation of 0.07 . This suggests a difference of about 1 to 2 percent of firms' size, with large variation between the two measures.³⁵

The results reveal that the null hypotheses of $\alpha = 0$ ($p = 0.609$) and $\beta = 1$ ($p = 0.240$) cannot be rejected, with point estimates of $\alpha = 0.01$ and $\beta = 0.8$.³⁶ Overall, although the adjustment procedure does not use data about the timing and amounts of all of the firms' transactions over the life of the firms until the reporting date (which are needed for complete inflation adjustment), the findings reveal that the algorithm provides a reasonable and unbiased proxy for the effects of inflation.

REFERENCES

- Aboody, D., M. E. Barth, and R. Kasznik. 1999. Revaluations of fixed assets and future firm performance: Evidence from the UK. *Journal of Accounting and Economics* 26 (1–3): 149–178.
- Ball, R. 1994. The development, accomplishments and limitations of the theory of stock market efficiency. *Managerial Finance* 20 (2–3): 3–48.
- , and E. Bartov. 1996. How naive is the stock market's use of earnings information? *Journal of Accounting and Economics* 21 (3): 319–337.
- Bar-Yosef, S., and B. Lev. 1983. Historical cost earnings versus inflation-adjusted earnings in the dividend decision. *Financial Analysts Journal* 39 (2): 41–50.
- Barth, M. E. 1994. Fair value accounting: Evidence from investment securities and the market valuation of banks. *The Accounting Review* 69 (1): 1–25.
- , W. H. Beaver, J. R. M. Hand, and W. R. Landsman. 1999. Accruals, cash flows, and equity values. *Review of Accounting Studies* 4 (3–4): 205–229.
- , D. P. Cram, and K. K. Nelson. 2001. Accruals and the prediction of future cash flows. *The Accounting Review* 76 (1): 27–58.
- , Y. Konchitchki, and W. R. Landsman. 2010. Cost of capital and earnings transparency. Working paper, Stanford University.
- Basu, S., S. Markov, and L. Shivakumar. 2006. Systematic errors in analysts' forecasts: Evidence from analysts' use of inflation information. Working paper, Emory University and London Business School.
- Beaver, W. H. 1979. Accounting for inflation in an efficient market. In *The Impact of Inflation on Accounting: A Global View*. *The International Journal of Accounting* 21–42. London, U.K.: Elsevier.
- . 1981. Market efficiency. *The Accounting Review* 56 (1): 23–37.
- , A. A. Christie, and P. A. Griffin. 1980. The information content of SEC Accounting Series Release No. 190. *Journal of Accounting and Economics* 2 (2): 127–157.
- , P. A. Griffin, and W. R. Landsman. 1983. How well does replacement cost income explain stock return? *Financial Analysts Journal* 39 (2): 26–30.

³⁵ The validation analysis requires that several obstacles be overcome. First, because the requisite Israeli data are not available in organized format, I hand-collect firms' annual data, as described above. Second, because Israeli GAAP requires footnote disclosure of selected nominal data, considerably more data are reported on an inflation-adjusted basis. Thus, I use an inverted algorithm that maps from inflation-adjusted to nominal amounts, and use as input the Israeli inflation-adjusted data. Third, I create a translation dictionary that classifies different accounting terms with the same content under a specific term and matches each Israeli data item to the equivalent Compustat item. This procedure results in Israeli firm-year observations with a format similar to that of U.S. companies in Compustat.

³⁶ I conduct further checks on the algorithm's accuracy. First, I form a statistic based on the mean difference between reported nominal earnings and earnings obtained from the algorithm, denoted as μ_x , and test $H_0: \mu_x = 0$ against $H_1: \mu_x \neq 0$. The results show that the null cannot be rejected ($p = 0.744$), which suggests the algorithm provides a reasonable estimate of the effects of inflation. Second, the algorithm uses computations that interact accounting items with monthly CPI values. To investigate whether these computations introduce measurement error, I derive *IAEarnings* after injecting a constant zero inflation rate into the system. This check results in *IAEarnings* being equal to *NominalEarnings*, consistent with zero inflation and zero measurement error from CPI computations. Third, I derive *NominalEarnings* using the algorithm and compare it to the Compustat amount. The results show the same earnings amount in all observations except those with missing values because of unavailable data.

- , M. McNichols, and R. Price. 2007. Delisting returns and their effect on accounting-based market anomalies. *Journal of Accounting and Economics* 43 (2–3): 341–368.
- , and W. R. Landsman. 1983. *Incremental Information Content of Statement 33 Disclosures*. FASB Research Report. Norwalk, CT: FASB.
- Bernard, V. L. 1986. Unanticipated inflation and the value of the firm. *Journal of Financial Economics* 15 (3): 285–321.
- , and C. Hayn. 1986. Inflation and the distribution of the corporate income tax burden. *National Tax Journal* 39 (2): 171–187.
- , and J. K. Thomas. 1990. Evidence that stock prices do not fully reflect the implications of current earnings for future earnings. *Journal of Accounting and Economics* 13 (4): 305–340.
- Campbell, J. Y., and T. Vuolteenaho. 2004. Inflation illusion and stock prices. *The American Economic Review* 94 (2): 19–23.
- Carhart, M. M. 1997. On persistence in mutual fund performance. *The Journal of Finance* 52 (1): 57–82.
- Chordia, T., and L. Shivakumar. 2005. Inflation-illusion and post-earnings-announcement drift. *Journal of Accounting Research* 43 (4): 521–556.
- Core, J. E., W. R. Guay, and R. Verdi. 2008. Is accruals quality a priced risk factor? *Journal of Accounting and Economics* 46 (1): 2–22.
- Davidson, S., C. P. Stickney, and R. L. Weil. 1976. *Inflation Accounting: A Guide for the Accountant and the Financial Analyst*. New York, NY: McGraw-Hill.
- Davis-Friday, P. Y., and E. A. Gordon. 2005. Relative valuation roles of equity book value, net income, and cash flows during a macroeconomic shock: the case of Mexico and the 1994 currency crisis. *Journal of International Accounting Research* 4 (1): 1–21.
- Dechow, P. M., R. G. Sloan, and A. Sweeney. 1995. Detecting earnings management. *The Accounting Review* 70 (2): 193–225.
- Easman, W., A. Falkenstein, and R. Weil. 1979. The correlation between sustainable income and stock returns. *Financial Analysts Journal* 35 (5): 44–48.
- Fama, E. F., and K. R. French. 1992. The cross-section of expected stock returns. *The Journal of Finance* 47 (2): 427–465.
- , and ———. 1993. Common risk factors in the returns of stocks and bonds. *Journal of Financial Economics* 33 (1): 3–56.
- , and J. D. MacBeth. 1973. Risk, return, and equilibrium: Empirical tests. *The Journal of Political Economy* 81 (3): 607–636.
- , and G. W. Schwert. 1977. Asset returns and inflation. *Journal of Financial Economics* 5 (2): 115–146.
- Financial Accounting Standards Board (FASB). 1978. *Objectives of Financial Reporting by Business Enterprises*. Statement of Financial Accounting Concepts No. 1. Norwalk, CT: FASB.
- . 1984. *Recognition and Measurement in Financial Statements of Business Enterprises*. Statement of Financial Accounting Concepts No. 5. Norwalk, CT: FASB.
- Gheyara, K., and J. Boatsman. 1980. Market reaction to the 1976 replacement cost disclosures. *Journal of Accounting and Economics* 2 (2): 107–125.
- Gordon, E. A. 2001. Accounting for changing prices: The value relevance of historical cost, price level, and replacement cost accounting in Mexico. *Journal of Accounting Research* 39 (1): 177–200.
- Gow, I. D., G. Ormazabal, and D. J. Taylor. 2010. Correcting for cross-sectional and time-series dependence in accounting research. *The Accounting Review* 85 (2): 483–512.
- Grossman, S. J., and J. E. Stiglitz. 1980. On the impossibility of informationally efficient markets. *The American Economic Review* 70 (3): 393–408.
- Hackel, K. S., J. Livnat, and A. Rai. 2000. A free cash flow investment anomaly. *Journal of Accounting, Auditing & Finance* 15 (1): 1–24.
- Hirshleifer, D., K. Hou, S. H. Teoh, and Y. Zhang. 2004. Do investors overvalue firms with bloated balance sheets? *Journal of Accounting and Economics* 38: 297–331.
- Kothari, S. P., and J. Shanken. 2004. Asset allocation with inflation-protected bonds. *Financial Analysts Journal* 60 (1): 54–70.
- Lettau, M., and S. Ludvigson. 2001. Resurrecting the (C)CAPM: A cross-sectional test when risk premia are time-varying. *The Journal of Political Economy* 109 (6): 1238–1287.

- Mishkin, F. S. 1983. *A Rational Expectations Approach to Macroeconomics: Testing Policy Ineffectiveness and Efficient-Market Models*. Chicago, IL: University of Chicago Press for the National Bureau of Economic Research.
- Modigliani, F., and R. A. Cohn. 1979. Inflation, rational valuation, and the market. *Financial Analysts Journal* 35 (2): 24–44.
- Petersen, M. A. 2009. Estimating standard errors in finance panel data sets: Comparing approaches. *Review of Financial Studies* 22 (1): 435–480.
- Petkova, R. 2006. Do the Fama-French factors proxy for innovations in predictive variables? *The Journal of Finance* 61 (2): 581–612.
- Phelps, E. S. 1967. Phillips curves, expectations of inflation, and optimal unemployment over time. *Economica* 34 (135): 254–282.
- Ritter, J. R., and R. S. Warr. 2002. The decline of inflation and the bull market of 1982–1999. *Journal of Financial and Quantitative Analysis* 37 (1): 29–61.
- Ro, B. T. 1980. The adjustment of security returns to the disclosure of replacement cost accounting information. *Journal of Accounting and Economics* 2 (2): 159–189.
- Shumway, T., and V. A. Warther. 1999. The delisting bias in CRSP's Nasdaq data and its implications for the size effect. *The Journal of Finance* 54 (6): 2361–2379.
- Sloan, R. G. 1996. Do stock prices fully reflect information in accruals and cash flows about future earnings? *The Accounting Review* 71 (3): 289–315.
- Summers, L. H. 1981. Inflation and the valuation of corporate equities. Working Paper No. 824. Washington, D.C.: National Bureau of Economics Research.
- Unites States Regulations. 2003. *Property, Plant, and Equipment Departmental Regulation*. Washington, D.C.: Office of the Chief Financial Officer.
- Watts, R. L., and J. L. Zimmerman. 1980. On the irrelevance of replacement cost disclosures for security prices. *Journal of Accounting and Economics* 2 (2): 95–106.

The Effect of Annual Report Readability on Analyst Following and the Properties of Their Earnings Forecasts

Reuven Lehavy
Feng Li
Kenneth Merkley
University of Michigan



ABSTRACT: This study examines the effect of the readability of firms' written communication on the behavior of sell-side financial analysts. Using a measure of the readability of corporate 10-K filings, we document that analyst following, the amount of effort incurred to generate their reports, and the informativeness of their reports are greater for firms with less readable 10-Ks. Additionally, we find that less readable 10-Ks are associated with greater dispersion, lower accuracy, and greater overall uncertainty in analyst earnings forecasts. Overall, our results are consistent with the prediction of an increasing demand for analyst services for firms with less readable communication and a greater collective effort by analysts for firms with less readable disclosures. Our results contribute to the understanding of the role of analysts as information intermediaries for investors and the effect of the complexity of written financial communication on the usefulness of this information.

Keywords: *financial analysts; readability; Gunning Fog Index; analyst following.*

Data Availability: *The data used in this study are available from the sources indicated in the text.*

I. INTRODUCTION

Over the past two decades, changes in financial and reporting regulations (e.g., changes in segment disclosures, employee stock option reporting, and Sarbanes-Oxley disclosures) have significantly increased the amount of required disclosures by firms to external users. In addition, technological advancement and new developments in financial engineering have made it more challenging for firms to communicate information about the underlying fundamentals of their businesses in a clear and informative manner. The increase in the amount of required dis-

We thank Beverly Walther (editor), Steven Kachelmeier (senior editor), and two anonymous reviewers. We also thank workshop participants from The University of Chicago, University of Michigan, University of Minnesota, Erasmus University, University of Melbourne, Tel Aviv University, the 2010 AAA Annual Meeting, the 2010 EEA Annual Congress, and a joint presentation of HEC-Paris and INSEAD for their helpful comments and suggestions.

Editor's note: Accepted by Beverly Walther.

Submitted: December 2009
Accepted: September 2010
Published Online: April 2011

closure accompanied by the challenging task of communicating increasingly complex business transactions to investors has led to concerns about the effectiveness of management communication and the ability of interested users to make informed decisions based on this information. As a result, both the SEC and the popular press have routinely criticized firms for the complexity of the language in their 10-K filings (Schroeder 2002). These groups have expressed concerns about the ability of financial statement users, especially small investors, to understand the complicated writing found in firms' financial documents (SEC 1998; Schroeder 2002; Cox 2007).

Given the increasing complexity of firm disclosures and the related concerns about their usability, a natural question arises as to the role of financial analysts in intermediating such information to investors. Specifically, do financial analysts use their expertise to examine this complex communication and provide useful information to financial statement users, or do they prefer to focus their efforts on firms with less complex communication? We attempt to answer this question by examining how the behavior of financial analysts relates to the readability of firms' written communication.

Prior literature examines the relation between the properties of firms' financial disclosures and the behavior of security analysts. These properties include the informativeness of disclosures (Lang and Lundholm 1996; Healy et al. 1999), the use of segment disclosures (Botosan and Harris 2000), and the effect of intangible assets (Barth et al. 2001). These studies generally document that firms with better disclosure quality tend to attract greater analyst following. Other studies examine the effect of certain financial reporting items on the properties of analyst earnings forecasts. These studies find that analyst forecasts are influenced by firms' accounting choices (Hopkins et al. 2000; Bradshaw et al. 2008), changes in the tax law (Plumlee 2003), the clarity of the income effects of specific financial items (Hirst and Hopkins 1998; Hirst et al. 2003), and international diversification (Duru and Reeb 2002).

While each of these studies contributes to our understanding of the effect of firms' financial reporting and disclosure choices on analyst behavior, they generally focus on the effect of a single financial statement item (e.g., interest rates, taxes) or disclosure (e.g., segment reporting).¹ This approach, however, does not explicitly incorporate the potential effect of the properties of other financial statement items or disclosures on analyst following or the properties of their earnings forecasts. More important, this approach does not address the potential impact on analyst behavior of the *overall* complexity of firms' financial communication, that is, the cost of processing and interpreting the entirety of firms' disclosures. This potential limitation is particularly important, given the broad nature of corporate disclosures such as 10-K filings that describe a significant number of interrelated financial items.

We attempt to address this issue by examining the relation between a comprehensive measure of the overall readability of corporate 10-K filings and analyst behavior. This measure, known as the Gunning Fog Index (hereafter, Fog Index), incorporates the number of words per sentence and the number of complex words in a document to derive a measure of the readability or syntactic complexity of firms' 10-K filings (Jones and Shoemaker 1994; Courtis 1995; Li 2008).² The Fog Index has been widely used in social science research for several decades to examine the relation

¹ One notable exception is Lang and Lundholm (1996) who use AIMR scores as a measure of overall disclosure quality. The potential limitations of this measure, however, are that it is based on a subjective survey of analysts, limited to a subset of large firms, and no longer available after 1995.

² Specifically, we focus on the readability of firms' written communication as a measure of the syntactic complexity of such communication. Stimulated by Chomsky's (1965) seminal generative grammar, there has been a great deal of theoretical and empirical work on what constitutes syntactic complexity and how it functions as a determinant of both sentence comprehension (Gibson 1998) and sentence production (Thompson and Farooqi-Shah 2002). According to Stone et al. (2005, 341–343), sentence complexity is affected by the features of open-class words (nouns and verbs) and their relationships and the number and type of syntactic operations.

between the readability of written information and various decisions or outcomes.³ Further, recent studies use the Fog Index to examine the readability of annual reports in connection with earnings persistence (Li 2008), timely price adjustment (Callen et al. 2009), and investment efficiency (Biddle et al. 2009).

The Fog Index offers several important advantages. First, it allows us to study a large and diverse group of firms. Second, it is an objective measure, not based on analyst surveys or opinion, and can be calculated for any narrative disclosure. Finally, it allows us not only to capture the effect of the readability of a variety of financial items, but, more important, to directly examine the overall syntactic complexity of firms' written public communication, over and above its specific content.⁴

We focus on the readability of the 10-K filing for several reasons. First, 10-K filings are required for all publicly traded companies and are frequently cited as an input in the decisions of investors and financial analysts (Previts et al. 1994; Rogers and Grant 1997). Second, both the SEC and popular press have criticized firms for the complexity of their language in these filings and the effect of this complexity on investors (Schroeder 2002). You and Zhang (2009) provide evidence that suggests that investors underreact to the information provided in 10-K filings and that this effect is stronger for firms with more complex 10-K reports. Finally, the 10-K filing contains a significant amount of written communication or narrative to use in interpreting its readability. These reasons make the 10-K filing an interesting setting to examine the effects of the readability of firm communication on the activities of financial analysts.

We test the relation between the readability of 10-K filings and several measures of analyst behavior, including analyst following, analyst forecast revision response time, the information content of analysts' reports, and the properties of analyst earnings forecasts. We begin our analysis by examining the relation between analyst following and 10-K filing readability. We argue that, on the one hand, lower readability of firm financial disclosures increases the cost of processing the information in these disclosures and therefore will increase the demand for analyst services. If the cost of obtaining analyst reports is less than the information-processing costs of firm disclosures, then analyst following should be greater for firms with less readable 10-K filings. On the other hand, less readable disclosure could also increase the costs of analyst coverage. That is, analysts could bear greater information-processing costs and higher private search costs, which could lead to less accurate forecasts. If these costs are significant, then analyst following should be lower for firms with less readable disclosures.

Consistent with the prediction of an increasing demand for analyst services for firms with less readable 10-K filings, we document a positive and significant association between a firm's 10-K Fog Index and the number of analysts who cover the firm, even after controlling for other factors related to analyst coverage. We interpret this result as being consistent with a greater collective effort by analysts for firms with less readable disclosures.

Next, we examine whether individual analyst effort is associated with the readability of the 10-K filing. We measure individual analyst effort as the length of time required for analysts to issue their first forecast revision following the 10-K filing. If analysts bear costs in following firms

³ The Fog Index has been used in a variety of applications including medical error, consumer drug use, consumer warranties, mutual fund prospectuses, jury instructions, and academic research prestige. For specific examples, see Charrow and Charrow (1979), Armstrong (1980), Shuptrine and Moore (1980), Ott and Hardie (1997), Gazmararian et al. (1999), Koo et al. (2003), and Lee et al. (2006).

⁴ In a recent speech, former SEC Chairman Christopher Cox specifically identifies readability measures as a tool to examine communication complexity. He states, "Just as the Black-Scholes model is commonplace when it comes to compliance with the stock option compensation rules, we may soon be looking to the Gunning-Fog and Flesch-Kincaid models to judge the level of compliance with the plain English rules" (Cox 2007). Similarly, Core (2001) proposes the use of computational linguistics methods in accounting to measure corporate disclosure quality.

with less readable disclosures, then it should take them more time, on average, to issue reports following less readable disclosures. Using a variety of empirical approaches, we find that analysts who cover firms with less readable 10-K filings take a longer time to issue their first report following the 10-K filing. This evidence is consistent with the argument that analysts exert more effort to follow firms with less readable disclosures.

We then turn to an examination of the relation between 10-K readability and the information content of analyst reports. Similar to Frankel et al. (2006), we measure information content as the proportion of a firm's stock returns related to analyst forecast revisions to the total firm's stock return during the time period between the 10-K filing and the subsequent fiscal year-end. We predict that, if investors place greater reliance on analyst reports issued for firms with less readable 10-K filings (potentially due to the higher costs associated with processing and interpreting a complex report), then the information content of analyst reports, or the proportion of firm returns associated with them, should be higher for firms with less readable disclosures. Consistent with this prediction, we find evidence that the proportion of firm returns associated with analysts' reports is higher for firms with less readable 10-K reports. This evidence suggests that investors find analysts' reports more informative for firms with less readable corporate disclosures.

Finally, we examine how 10-K filing readability relates to commonly studied properties of analyst earnings forecasts: dispersion, accuracy, and two measures of earnings consensus forecast uncertainty (overall and common uncertainty) based on Barron et al. (1998). Prior literature finds that the 10-K filing represents a major source of information used by analysts (Previts et al. 1994; Rogers and Grant 1997). However, studies also find that the complexity of items in the 10-K filing affects analysts' use of this information. For example, Plumlee (2003) finds that analyst forecasts are less accurate if they are associated with complex changes in the tax law. Bradshaw et al. (2008) find that differences in accounting choice negatively affect forecast accuracy and increase dispersion. Our tests relate to these findings, because we examine the effect of complexity through the ability of analysts to incorporate less readable financial information into their forecasts. If lower readability increases the costs of processing and interpreting firm disclosures, then it could limit the ability of analysts to correctly incorporate all of the pertinent information and could result in disagreement or ambiguity. We predict and find that analyst earnings forecasts for firms with less readable 10-K reports have greater dispersion, are less accurate, and are associated with greater overall analyst uncertainty. We also find that analyst common uncertainty is increasing in 10-K Fog. These results suggest that analyst forecasts are affected by 10-K readability and provide indirect evidence to support the notion that the information contained in the 10-K filing is used by analysts.

This study contributes to the literature in the following ways. First, our overall findings that lower readability is associated with greater levels of analyst coverage, effort, and information content, but with lower accuracy, higher dispersion, and greater uncertainty of their earnings forecasts complements and contributes to the literature on how analysts respond to firms' disclosure. While prior studies generally focus on the effect of the complexity of specific attributes of firms' disclosure on analyst behavior, we provide evidence that suggests that analyst behavior is associated with overall disclosure readability.⁵ While the collection of the evidence in this study is consistent with this inference, the evidence obtained from each individual test is subject to limitations common to our research design.

⁵ We note that covering firms with less readable communication is likely to result in a trade-off. Investors are more likely to find analyst coverage useful for firms with less readable communication; however, earnings forecasts for these firms will potentially be less accurate. Our findings suggest that the benefits of coverage outweigh the costs, because we document that both greater coverage and less accurate forecasts are associated with less readable communication (see also Li et al. 2009). This is consistent with the notion that investors value more than just forecast accuracy (Schipper 1991; Clement and Tse 2003).

Second, our findings relate to the debate about the intended audience of public corporate filings. Since its organization, the SEC has made consistent efforts to encourage firms to make their regulatory filings accessible to the average or “lay” investor (Firtel 1999).⁶ The most recent of these efforts is the plain English disclosure rules adopted by the SEC on January 22, 1998. Our evidence, that less readable corporate disclosures are associated with greater analyst following and more informative analyst reports, suggests that analysts could serve as an available information intermediary for investors and relates to the SEC’s concerns about the accessibility of these reports.

Section II discusses the role of the readability of firm communication and the hypotheses development. Section III describes the data and sample, and section IV presents our empirical evidence. Section V concludes.

II. HYPOTHESES DEVELOPMENT

Less readable communication is more difficult to interpret and process by investors because it requires that investors devote more time and effort to identify and extract relevant information (Bloomfield 2002). In this study we focus on the written complexity of firm disclosures, as measured by readability rather than content. We interpret disclosure readability to be a measure of the costs incurred by users to process and interpret a firm’s written communication after controlling for the operational complexity of the business.⁷ Firms with similar operations provide disclosures with varying levels of readability. For example, Berkshire Hathaway and AIG both provide a section about their reinsurance businesses in the management’s discussion and analysis section of their 10-K filings for fiscal year 2003. While the subject matter and underlying business complexity of the reinsurance business sections are very similar across the two firms, Berkshire Hathaway’s explanation is more readable as evidenced by a lower Fog Index.⁸ This is not surprising, insofar as Berkshire Hathaway CEO Warren Buffett is a strong proponent of more readable communication.

We begin our analysis by examining the effect of disclosure readability on analyst following. We assume that users of financial information have different abilities to process complex communication (Indjejikian 1991; Ball 1992). These differences provide opportunities for information intermediaries, such as financial analysts, to profit from their private analysis of firms by selling their opinions to users with greater information-processing costs (Schipper 1991). These profitable opportunities are arguably greater for firms with less readable disclosures because of the greater cost to users of processing the firms’ information. As such, if analysts respond to this increased demand for their services, then we expect analyst following to be greater for firms with less readable disclosures.

⁶ Some academic scholars and practitioners have argued that the primary users of this information should be market professionals such as analysts and not the “lay” investor (Kripke 1970; Schipper 1991). This is because nonprofessionals might not possess the skill or expertise to read and understand the complex financial information contained in disclosure documents and, therefore, any effort to gear disclosure toward the layperson is a waste of time and money (Kripke 1970). Further, critics argue that the SEC’s efforts to appeal to the average investor are not only inefficient, but also are hazardous to the disclosure regime because simplified and concise disclosure often leaves out many issues that are potentially valuable to the professional and leads to potential legal liability.

⁷ It is important to note that more readable communication can come at a cost. Legal scholars point out that “it is much harder to simplify than to complicate” (Kimble 1994, 53) and it requires significant skill, work, and time to compose documents in plain language. Additionally, some critics argue that plain English communication increases the risk of litigation (Kripke 1973).

⁸ The Fog Index for AIG’s reinsurance business section of the MD&A is 18.51, while Berkshire Hathaway’s is 17.23. The difference between these scores (1.28) can be interpreted as the number of additional years of formal education required to understand the text on a first reading. Further details regarding the Fog Index are provided in Section III.

We note, however, that analysts also face a variety of costs in covering firms with less readable communication. First, analysts bear the direct costs of processing the information provided by management. Second, less readable communication can result in confusion, requiring analysts to incur greater private search costs as they obtain additional information to evaluate and interpret management's communication. Third, greater syntactic complexity could lead to inaccurate forecasts and recommendations that adversely affect analysts' careers (e.g., Mikhail et al. 1999; Plumlee 2003; Hong and Kubik 2003).⁹ Finally, Li (2008) finds that firms attempt to obfuscate bad news by increasing the complexity of their communication, and Lang et al. (2004) find that analysts are less likely to cover firms with incentives to withhold or manipulate information. These potential manipulations along with the costs to analysts of covering firms with less readable communication discourage analyst following. Because there are both potential positive and negative consequences of syntactic complexity on analyst following, we test the null hypothesis that syntactic complexity has no association with analyst following.

Given the difficulty of following firms with less readable disclosures, analysts who choose to follow these firms likely exert greater effort to do so. One way of measuring the effort that analysts exert in following firms with less readable disclosures is to examine the average time from the firm's 10-K filing to the analyst's first report subsequent to the filing. If lower readability results in higher processing costs to analysts, then firms with less readable communication will have greater average response time from the analysts that cover them as compared to the analysts covering firms with more readable communication. Similarly, analysts who cover a portfolio of firms that have less readable communication should take longer, on average, to issue their reports than other analysts. We test the hypothesis that analysts exert greater individual effort to cover firms with less readable disclosures.

Readability could also affect the properties of analyst earnings forecasts. Specifically, we examine its effect on the information content, accuracy, and dispersion of analyst forecasts and the uncertainty in analysts' information environment. Analyst reports for firms with less readable communication likely provide information that is more useful to investors due to greater costs in processing public information. Analyst reports for firms with less readable communication would also be more informative if analysts choose to acquire and incorporate more private information due to the difficulties involved in interpreting less readable disclosures. Accordingly, we predict that analyst reports for firms with less readable communication will have greater information content, on average, than firms with more readable communication.

Finally, because analysts who follow firms with less readable communication bear the costs of processing and interpreting such disclosures, syntactic complexity could also affect the dispersion and accuracy of analyst earnings forecasts as well as the degree of the overall and common uncertainty embedded in analyst forecasts. If less readable communication increases analysts' cost to process and interpret the information, then it is likely to lead to a more diverse set of interpretations about firm disclosures, resulting in higher analyst forecast dispersion. In addition, if less readable communication makes it more difficult to forecast earnings, then the accuracy of the analyst consensus forecast will be lower for firms with less readable communication. However, this difference in accuracy can be fully or partially offset by greater analyst effort in response to greater syntactic complexity. Accordingly, we predict that analyst forecasts will have greater dispersion and be less accurate for firms with less readable communication.

⁹ Li et al. (2009) find that once analysts become all-stars they are more willing to cover firms with greater potential earnings management because investors could find their coverage more valuable for these firms. This suggests that coverage decisions can be related to favorable career outcomes; however, it is not clear how analysts assess the trade-offs between the demand for investment information and the cost of potentially less accurate forecasts.

Barron et al. (1998) point out that analyst earnings forecast dispersion and accuracy do not directly capture the theoretical properties of analysts' information environment such as uncertainty because these measures relate in different ways to the idiosyncratic and the common components of error in analysts' forecasts. They view idiosyncratic uncertainty as uncertainty in private information that affects each analyst differently, and common uncertainty as uncertainty in information that is common to all analysts. Accordingly, lower readability can be associated with higher dispersion and lower accuracy in earnings forecasts for reasons related to either or both common and idiosyncratic uncertainty in the analysts' information environment.

To further examine the association between readability and properties of analyst earnings forecasts, we follow the model derived in Barron et al. (1998) and construct empirical measures of the overall uncertainty (modeled as the sum of the idiosyncratic and the common uncertainty) and the common uncertainty in analyst forecasts. These measures combine the accuracy, the dispersion, and the number of analyst forecasts, allowing us to examine more directly how readability relates to analysts' information environment uncertainty. We predict that analyst forecasts for firms with less readable reports will be associated with greater overall uncertainty. The effect of readability on common uncertainty (i.e., how much the average belief reflects common versus private information) is less clear because analysts can make trade-offs between public and private information sources based on the relative precision and importance of these sources.

III. SAMPLE AND VARIABLE DEFINITIONS

Sample Selection

Our initial sample is based on the intersection of firm/years available on the Compustat Fundamental Annual table and the SEC's EDGAR filings database for fiscal years 1995–2006. These databases are joined based on Compustat GVKEY and the SEC's Central Index Key (CIK). Firms without matches are dropped from the sample. For each firm-year observation, we download the corresponding 10-K filing. Filings with less than 3,000 words or 100 lines are dropped to ensure that a complete filing is examined and that no errors were made in the filing transmission. This procedure results in 57,642 observations. We obtain stock return data from CRSP, analyst data from I/B/E/S, institutional holdings data from Thomson Reuters (CDA/Spectrum), and information on management earnings guidance from the First Call Company Issued Guidelines database. This procedure yields a sample of 33,704 observations.

Disclosure Readability

Similar to Li (2008), we measure the readability of 10-K filings using the Fog Index (*FOG*). This index, developed in the computational linguistics literature, captures the written complexity of a document as a function of the number of syllables per word and the number of words per sentence. Specifically, we calculate the readability of the 10-K report for firm *i* in year *t* as follows:

$$FOG_{i,t} = (\text{average words per sentence} + \text{percent of complex words}) \times 0.4 \quad (1)$$

where a complex word is defined as one with three or more syllables. The index is interpreted as the number of years of formal education required for a person of average intelligence to read the document once and understand it.¹⁰ It is important to note that the Fog Index is a measuring tool,

¹⁰ Before computing the Fog Index, we remove all tables, tabulated text, and financial statements from the 10-K. The constant of 0.4 found in Equation (1) was chosen by Robert Gunning based on the scores of a set of literary benchmarks in order that this specific interpretation could be made. Our inferences remain unchanged using alternative measures of readability such as the Kincaide grade level formula, the Flesch Reading Ease Index, and 10-K length. In addition, our main inferences are similar after including in the analysis the length of the 10-K filing as a control variable for the amount of disclosure provided by the firm (Leuz and Schrand 2009).

not a rule or formula for good writing; as stated by Gunning (1969, 12), "Nonsense written simply is still nonsense." It predicts the readability of a document, but does not provide information about whether the writing is interesting or informative. Despite these limitations, it is objective and simple to calculate. It allows us to study the disclosure characteristics of a large and diverse group of firms and does not depend on analyst surveys or opinions. It also provides us with a comprehensive measure of the overall syntactic complexity of 10-K filings as opposed to the complexity of individual financial items.¹¹

Table 1 provides descriptive statistics on the Fog Index for firms in the intersection of Compustat and EDGAR. As can be seen in Panel A, the overall mean and median of the Fog Index are 19.52 and 19.38, respectively. As a comparison to help validate our measure, it is interesting to note that the average Fog Index for FASB Statement Nos. 1–122 is 22, the CPA Exams is 16, the CMA Exams is 17, *Wall Street Journal* is 12, and *Reader's Digest* is 8 (Phillips et al. 2007; Cox 2007). The relatively high mean and median Fog Index in our sample is consistent with concerns that financial reports are written in complex language.

While the variation in the mean Fog Index over the sample period is modest, there is large variation within each year. For example, over the total sample years, the interquartile range for the Fog Index is from 1.58 to 1.88 (about 9 percent of the mean). There is also significant variation in the Fog Index within industries, despite similarities in the underlying business complexity within each industry. Table 1, Panel B provides examples of specific industries that have high and low levels of the Fog Index (industries are classified using the Fama and French 48-industry classification). The healthcare, insurance, trading (e.g., security brokers, investment offices, etc.), utilities, and telecommunications industries comprise the group with the highest Fog Index. The fact that these industries have a low level of disclosure readability is not surprising because they are characterized by complex contracts and business models that are difficult to communicate. However, there is also significant variation within each industry. Industries with a low Fog Index comprise precious metals, shipping containers, food products, agriculture, and defense industries. While the precious metals and shipping container industries have small within-industry variation, most industries have an interquartile range of the Fog Index above 1.60. The relative rankings of these industries also help to validate the ability of the Fog Index to measure the readability of annual reports.

Variable Definitions

Our tests focus on examining the relation between the Fog Index and analyst following, their forecast revision response time, the information content of their reports, and the properties of their earnings forecasts. A description of the construction of these variables as well as our control variables follows.

Analyst Following

Similar to prior research (O'Brien and Bhushan 1990; Brennan and Subrahmanyam 1995), we define analyst following as the number of analysts (*#ANALYSTS*) that comprise the first I/B/E/S consensus annual earnings forecast after the filing date of the 10-K report. We follow Bhushan (1989) and interpret this measure as a proxy for the collective effort of the financial analyst community in the analysis of an individual firm.¹² Because some firms are not covered by I/B/E/S,

¹¹ Prior literature used the Fog Index and other similarly constructed measures to examine the readability of the overall annual report (Jones and Shoemaker 1994), management's discussion and analysis (Schroeder and Gibson 1990), and the notes to the financial statements (Smith and Smith 1971; Healy 1977).

¹² Similar to Bhushan (1989), we acknowledge that our proxy is not a perfect measure because it assumes homogeneity among analyst effort levels. For example, there are differences in individual analysts' effort levels based on differences in compensation, brokerage houses, etc.

TABLE 1
Descriptive Statistics for Disclosure Readability Using the Fog Index of the 10-K Filing

Panel A: Fog Index by Fiscal Year (All Compustat Firms)									
Year	n	Mean	Std. Dev.	Q1	Median	Q3	Interquartile		
1995	3,107	19.25	1.31	18.38	19.17	20.04	1.66		
1996	5,221	19.45	1.31	18.53	19.38	20.24	1.71		
1997	5,530	19.48	1.30	18.59	19.39	20.28	1.70		
1998	5,329	19.44	1.25	18.59	19.35	20.20	1.62		
1999	5,524	19.26	1.29	18.40	19.14	20.02	1.62		
2000	5,494	19.19	1.29	18.34	19.05	19.92	1.58		
2001	4,955	19.30	1.42	18.39	19.14	20.00	1.61		
2002	4,723	19.57	1.52	18.60	19.38	20.33	1.73		
2003	4,546	19.91	1.79	18.80	19.63	20.68	1.88		
2004	4,425	19.78	1.86	18.81	19.64	20.67	1.87		
2005	4,449	19.86	1.50	18.88	19.68	20.56	1.68		
2006	4,339	19.90	1.49	18.94	19.72	20.61	1.67		
All Years	57,642	19.52	1.47	18.58	19.38	20.29	1.70		

Panel B: Fog Index by Industry (All Compustat Firms)							
	n	Mean	Std. Dev.	Q1	Median	Q3	Interquartile
Industries with High Fog Index							
Healthcare	1,115	20.22	1.46	19.30	20.04	20.99	1.69
Insurance	2,143	20.16	1.29	19.32	20.10	20.89	1.57
Trading	3,279	19.88	1.67	18.78	19.75	20.75	1.97
Utilities	2,405	19.84	1.56	18.92	19.64	20.53	1.61
Telecommunications	2,021	19.80	1.45	18.89	19.59	20.45	1.56
Industries with Low Fog Index							
Precious Metals	135	18.43	0.94	17.75	18.38	19.12	1.37
Shipping Containers	159	18.64	1.25	17.87	18.43	19.20	1.33
Food Products	483	18.86	1.56	17.95	18.65	19.63	1.68
Agriculture	162	18.95	1.47	17.90	18.85	19.58	1.68
Defense	108	18.99	1.45	18.00	18.87	19.88	1.88

(continued on next page)

Panel A reports descriptive statistics on our measure of disclosure readability, the Fog Index, for observations available on the SEC's EDGAR and Compustat databases. The Fog Index is computed as: (average words per sentence + percent of complex words) $\times$ 0.4, using the text of the 10-K filings for fiscal years 1995–2006.

Panel B provides descriptive statistics for the five industries with the highest and lowest mean Fog Index. Industries are based on the Fama and French 48-industry classification.

we conduct tests for both the I/B/E/S sample and the full sample of firms, where missing coverage is coded as zero analyst coverage (Barth et al. 2001). Because these results are similar, we report our main results based on the I/B/E/S sample.

Analyst Forecast Revision Response Time

We define the analyst forecast revision response time as the time from the 10-K filing to the first annual or quarterly earnings forecast issued by each *individual* analyst following the firm. To ensure that we include only analysts who actively follow the firm, we require that each analyst issue a forecast in the 90 days prior to the 10-K filing and then another report within 90 days of the 10-K filing (this process eliminates 10,458 observations related to analysts who stopped coverage). Because earnings announcements can prompt analysts to issue reports, we exclude reports made after any earnings announcement that occurs after the 10-K filing, but before the end of the 90-day window. We then define analyst report duration as the length of time in days between the 10-K filing and the first report following the filing. In addition, we average individual analyst report duration at both the individual firm and analyst level for each year. We interpret this duration variable as a measure of the required amount of effort for an individual analyst to read, understand, and process the information contained in the 10-K filing and to issue an updated earnings forecast. As mentioned above, we expect analyst response time to be longer for less readable 10-K filings.

Information Content of Analyst Reports

Similar to Frankel et al. (2006), we measure the information content of analyst reports as the proportion of a firm's stock returns related to analyst forecast revisions to the total stock return during the time period between the 10-K filing and the subsequent fiscal year-end. This measure is constructed as the sum of the one-day, absolute size-adjusted returns on the analyst forecast revision day, divided by the sum of the one-day, absolute size-adjusted returns over the entire window.¹³ We exclude analyst reports that coincide with earnings announcements. We treat multiple reports issued on the same day as a single report. Also, an observation must have a minimum of 90 trading days with available data to be considered in our tests. Similar to Frankel et al. (2006), we interpret this measure as the percentage of total firm information provided to investors that is related to analyst reports.

Properties of Analyst Earnings Forecasts

We define analyst forecast dispersion (*DISPERSION*) as the standard deviation of the individual analyst forecasts in the first analyst consensus annual earnings forecast issued after the 10-K filing for the fiscal period following the 10-K filing, scaled by share price 90 days before the consensus forecast date. Analyst forecast accuracy (*ACCURACY*) is computed as the squared difference between I/B/E/S reported earnings and the analyst consensus forecast, scaled by share price 90 days before the consensus forecast date.¹⁴ We define analyst overall uncertainty

¹³ Frankel et al. (2006) divide this measure by the number of forecast revision dates to obtain a measure of the average informativeness of an analyst report date. We omit this final step for two reasons. First, our variable of interest is the overall firm information that comes from analysts, not the average information content. Second, firms with less readable disclosures are likely to obtain greater amounts of information from analysts because analysts are more likely to revise their forecasts more often for such firms. We confirm this supposition in untabulated results. Thus, dividing by the number of forecast revisions would induce a negative correlation between the analyst informativeness measure and 10-K readability. Our results are robust to the inclusion of the number of analysts following the firm as a control for the expected number of analyst revisions.

¹⁴ Our inferences are not sensitive to scaling the measures of analyst forecast accuracy and dispersion by the absolute value of earnings or using the unscaled variables.

($UNCERTAINTY_{OVERALL}$) and common uncertainty ($UNCERTAINTY_{COMMON}$) using the following equations derived by Barron et al. (1998):¹⁵

$$UNCERTAINTY_{OVERALL} = \left(1 - \frac{1}{\#ANALYSTS} \right) * DISPERSION + ACCURACY \quad (2)$$

$$UNCERTAINTY_{COMMON} = \frac{ACCURACY - \frac{DISPERSION}{\#ANALYSTS}}{UNCERTAINTY_{OVERALL}} \quad (3)$$

Following Barron et al. (1998), we interpret overall uncertainty ($UNCERTAINTY_{OVERALL}$) as the sum of the idiosyncratic uncertainty (i.e., uncertainty associated with analysts' private information) and common uncertainty (i.e., uncertainty associated with information common to all analysts). $UNCERTAINTY_{COMMON}$ is calculated as the proportion of common uncertainty to overall uncertainty. Thus, we interpret $UNCERTAINTY_{COMMON}$ to be a measure of how much the average analyst's belief reflects common versus private information. As shown in Equations (2) and (3), as $DISPERSION$ approaches 0, $UNCERTAINTY_{OVERALL}$ approaches $ACCURACY$ and $UNCERTAINTY_{COMMON}$ approaches 1. This suggests that, when there is no disagreement among analysts (i.e., $DISPERSION$ equals 0), the total uncertainty is only associated with analysts' common information and all information impounded in analyst forecasts is public. Empirically, because the uncertainty measures emphasize information across analysts, we require that each firm have at least four analysts following it.

Control Variables

Our analysis controls for a variety of variables that have been shown by prior literature to be associated with firms' information environment and business complexity and therefore relate to analyst behavior. Prior work finds that firm size is the most important determinant of analyst following (Bhushan 1989; O'Brien and Bhushan 1990; Brennan and Hughes 1991; Lang and Lundholm 1996; Barth et al. 2001). These studies find that larger firms have greater analyst following and suggest that large firms have better information environments, potentially more complex operations, and greater demand for investment advice. We use the natural logarithm of market value as of the year ending prior to the 10-K filing ($LOGSIZE$) as a proxy for size.

Following Barth et al. (2001), we include controls for growth. High-growth firms tend to attract greater analyst following due to investor interest and the potential for future investment banking deals. Further, analysts likely find it more difficult to accurately forecast earnings for firms with high growth, leading to greater disagreement among analysts and less accurate forecasts. We define the variable $GROWTH$ as the compounded average growth rate in sales over the prior three to five fiscal years. Our inferences are robust to the inclusion of the natural logarithm of the book-to-market ratio as a proxy for growth.

Similar to Bradshaw et al. (2008), we include the natural logarithm of the number of business segments reported in the Compustat Segment File as a control for the underlying complexity of the firm. We also include a control for the level of institutional holdings, following the evidence in Bhushan (1989), Brennan and Subrahmanyam (1995), and Frankel et al. (2006). These studies find that institutional ownership is positively associated with analyst following and with the informa-

¹⁵ We thank the reviewer for suggesting that we incorporate the Barron et al. (1998) measures into our analysis. Note that the variables we term overall and common uncertainty are termed analyst uncertainty and consensus, respectively, in Barron et al. (1998). We use the term common uncertainty as opposed to consensus to avoid confusion regarding our use of analyst consensus forecasts.

tion content of their reports.¹⁶ Institutional ownership can also be associated with higher analyst forecast accuracy and lower dispersion because firms with high levels of institutional holdings tend to have better information environments. We define the variable *PINST* as the percentage of a firm's shares that are held by institutions from the 13F disclosures for the most recent quarter prior to the 10-K filing. Motivated by the evidence in Lang and Lundholm (1996), who document that analyst following increases with the quality of disclosures, we include the number of management earnings forecasts made during the prior year (*MFCOUNT*) as a proxy for firm discretionary disclosure (Nagar et al. 2003; Cotter et al. 2006), as well as the absolute value of the cumulative, market-adjusted return for the two-day window around the 10-K filing (*10-K NEWS*) as a proxy for the disclosure informativeness.¹⁷

Similar to Barth et al. (2001), we include variables to control for firms' information environment and business complexity, as these variables likely affect the properties of analyst earnings forecasts. Barth et al. (2001) examine the association between analysts' incentives to follow firms and the extent of their intangible assets. Because many intangible assets are generally not recognized and estimates of their fair value are not disclosed, Barth et al. (2001) argue that analysts have increased incentives to follow firms with greater intangibles due to increased demand from investors. They find that analyst following is greater for firms with larger research and development and advertising expenses. They interpret these results as evidence that analysts respond to the demands of investors for more information because of the difficulty in evaluating firms' intangible assets. Other studies examine the effect of intangibles on the properties of analyst forecasts. Barron et al. (2002) find that analyst uncertainty increases with the level of a firm's intangible assets. Gu and Wang (2005) find that analyst forecast errors are increasing in firm intangible intensity. Intangibles are important for our tests because they could be associated with 10-K readability. For example, it is likely more difficult to explain the operations of firms with high levels of research and development costs in a less complex manner. Similar to Barth et al. (2001), we define *R&D* as the ratio of research and development expense to operating expense and *ADV* as the ratio of advertising expense to operating expense.¹⁸

Finally, we include the standard deviation of firm monthly stock returns from the prior year (*STD_RET*) as a measure of information uncertainty.¹⁹ Bhushan (1989) suggests that private information is more valuable for firms with higher return volatility and thus positively related to the demand for analyst services. However, it is likely that analysts bear increased costs for following firms with higher return volatility. Additionally, most of our tests include industry fixed effects (based on the Fama and French classification) and year fixed effects to account for variation in analyst following across specific industries and over time. This approach is used to help control for variation in business complexity or information uncertainty that is driven by industry or time.

¹⁶ O'Brien and Bhushan (1990) suggest that analyst following and institutional ownership are potentially endogenous. Frankel et al. (2006) assume that the variables are exogenous. Similar to Barth et al. (2001), we examine our results with and without the inclusion of institutional ownership and also with the inclusion of lagged institutional ownership. The coefficient on *FOG* remains positive and significant under these specifications.

¹⁷ We thank a reviewer for pointing out the importance of controlling for the informativeness of the 10-K disclosure.

¹⁸ Our results are robust to the inclusion of the amount of recognized intangibles and depreciation expense, which are also examined by Barth et al. (2001).

¹⁹ While some of our other control variables (e.g., sales growth and R&D) are likely correlated with firms' information uncertainty, we include this measure as a more explicit, market-based measure of this construct. In addition, the inclusion of an alternative measure, earnings volatility, does not significantly affect our results. However, the inclusion of this measure reduces our sample by about 20 percent due to the need for a sufficient time-series of earnings.

IV. RESULTS

Summary Statistics

Table 2 presents summary statistics for the sample of 33,704 firm-year observations that remain after imposing the availability of data on CRSP and I/B/E/S. Similar to Table 1 and consistent with concerns raised by the SEC about disclosure complexity, the mean (median) *FOG* score is 19.53 (19.38). The standard deviation and interquartile range are 1.42 and 1.71, respectively, with an interquartile range from 18.58 to 20.29. The mean (median) number of analysts per firm-year observation is 6.14 (4.00). The mean (median) average analyst forecast revision response time is 17.71 (16.00) days at the firm-year level and 18.77 (17.50) days at the analyst-year level. This suggests that, on average, analysts do not respond immediately to 10-K filings due to the significant amount of information that must be processed and interpreted. The mean (median) analyst report information content (*AI*) is 0.130 (0.092). This indicates that, on average, almost 13 percent of the information reflected in stock returns during the period between the 10-K filing and the end of the fiscal period is derived from analyst reports. The mean (median) of forecast dispersion is 0.008 (0.003) and that of squared forecast error (*ACCURACY*) is 0.078 (0.002). The mean (median) of *UNCERTAINTY_{OVERALL}* and *UNCERTAINTY_{COMMON}* are 0.051 (0.005) and 0.328 (0.267), respectively. This suggests that about 33 percent of analyst uncertainty about future earnings following the 10-K filing is based on public information.

Table 2 also provides statistics on the control variables. The mean (median) size of our sample firms is \$2.1 billion (\$381 million), and mean (median) compound averages growth rate of sales (*GROWTH*) is 0.20 (0.11). The mean (median) number of business segment is 1.83 (1), and mean (median) percent of institutional ownership is 46 (45). The mean (median) number of management earnings forecasts is 1.44 (0) and the mean (median) *10-K NEWS* is 0.03 (0.02). The mean (median) ratio of research and development expense (*R&D*) and advertising (*ADV*) to operating expense are 0.08 (0.00) and 0.01 (0.00), respectively.²⁰ The mean (median) standard deviation of returns (*STD_RET*) is 0.14 (0.12).

Similar to Li (2008), we find in untabulated results that, while *FOG* is significantly correlated with many of the control variables, the extent of these correlations is relatively small for most variables. This is consistent with the conclusion that it is difficult to explain a significant proportion of the variation in *FOG* using firm characteristics. This evidence is important because in this study we assume that the Fog Index measures the readability of 10-K disclosures rather than their content. Consistent with this assertion, the correlation between *FOG* and the absolute value of the 10-K event return is small and insignificant.

As a caveat, we note that disclosure readability can be decomposed into innate and discretionary components. While we include a variety of variables to control for innate readability, analysts' information costs are based on total readability, so the relative magnitude of these components is less relevant. Our inferences are limited by the extent to which the Fog Index measures readability and is not confounded with other (uncontrolled) firm characteristics. It is also important to note that, for a specific variable to influence the readability measure, it must either increase the length of the average sentence in the 10-K document or increase the percentage of complex words. Simply increasing the length of the 10-K will not directly affect the measure. However, we examine the effect of disclosure length on our main results in sensitivity tests because length also can be a form of complexity (You and Zhang 2009; Loughran and McDonald 2010; Miller 2010).

²⁰ Note that the medians for the number of management earnings forecasts, research and development expense, and advertising expense are zero. This is primarily because we follow prior literature and code missing items to be zero.

TABLE 2
Sample Descriptive Statistics

Variable	n	Mean	Std. Dev.	Q1	Median	Q3
FOG	33,704	19.53	1.42	18.58	19.38	20.29
#ANALYSTS	33,704	6.14	5.91	2.00	4.00	8.00
RESPONSE _{Firm}	17,868	17.71	10.76	9.50	16.00	24.25
RESPONSE _{Analyst}	5,737	18.77	8.49	12.50	17.50	23.50
AI	30,716	0.130	0.139	0.037	0.092	0.177
DISPERSION	26,078	0.008	0.018	0.001	0.003	0.008
ACCURACY	29,055	0.078	0.447	0.000	0.002	0.015
UNCERTAINTY _{OVERALL}	17,241	0.051	0.280	0.002	0.005	0.016
UNCERTAINTY _{COMMON}	17,222	0.328	0.401	-0.038	0.267	0.702
SIZE	33,704	2100	5600	122	381	1300
GROWTH	33,704	0.20	0.34	0.04	0.11	0.24
SEGMENTS	33,704	1.83	1.32	1.00	1.00	3.00
PINST	33,704	45.98	26.78	22.95	45.31	67.60
MFCOUNT	33,704	1.44	2.69	0.00	0.00	2.00
10-K NEWS	33,704	0.03	0.05	0.01	0.02	0.04
ADV	33,704	0.01	0.03	0.00	0.00	0.01
R&D	33,704	0.08	0.17	0.00	0.00	0.08
STD_RET	33,704	0.14	0.10	0.07	0.12	0.18

This table reports descriptive statistics for observations that are available on the SEC's EDGAR, Compustat, CRSP, and I/B/E/S databases. The variables are pooled across fiscal years 1995–2006.

Variable Definitions:

- FOG = Fog Index of the 10-K filing calculated as (average words per sentence + percent of complex words) $\times$ 0.4;
- #ANALYSTS = number of analysts in the first consensus annual earnings forecast following the 10-K filing;
- RESPONSE_{Firm} = average number of days that it takes a firm's analysts to issue their first report following the 10-K filing;
- RESPONSE_{Analyst} = average number of days that it takes an analyst to issue his/her first report following the 10-K filing;
- AI = sum of the one-day, absolute size-adjusted returns from analyst reports between the 10-K filing and the next fiscal year-end divided by the sum of the one-day, absolute size-adjusted returns over the entire window;
- DISPERSION = standard deviation of the individual analyst forecasts in the first analyst consensus annual earnings forecast issued after the 10-K filing for the fiscal period following the 10-K filing, scaled by share price 90 days before the consensus forecast date;
- ACCURACY = squared difference between I/B/E/S reported earnings and the first analyst consensus annual earnings forecast issued after the 10-K filing for the fiscal period following the 10-K filing, scaled by share price 90 days before the consensus forecast date;
- UNCERTAINTY_{OVERALL} = sum of common and idiosyncratic uncertainty among analysts computed following Barron et al. (1998, equation 15), using the measures of accuracy, dispersion, and analyst following as previously defined in this table;
- UNCERTAINTY_{COMMON} = ratio of common uncertainty to total uncertainty among analysts computed following Barron et al. (1998, equation 16) using the measures of accuracy, dispersion, and analyst following as previously defined in this table;
- SIZE = number of shares outstanding (Compustat item CSHO) times the share prices at the most recent fiscal year-end (Compustat item PRCC_F);
- GROWTH = compound average growth rate of firm sales (Compustat item SALE) over the prior 3–5 years $(sales_{t-1} / sales_{t-(1+t)})^{1/t}$;
- SEGMENTS = number of reported business segments in the Compustat segment file for prior fiscal year;
- PINST = percentage of institutional ownership from the quarter prior to the 10-K filing;

(continued on next page)

TABLE 2 (continued)

<i>MFCOUNT</i>	= number of management earnings forecasts issued in the prior year;
<i>10-K NEWS</i>	= absolute value of the cumulative market-adjusted return for the 10-K filing event window [0,1];
<i>ADV</i>	= advertising expense (Compustat item XAD) as a percentage of operating expense (Compustat item XOPR) from the prior fiscal year;
<i>R&D</i>	= research and development expense (Compustat item XRD) as a percentage of operating expense (Compustat item XOPR) from the prior fiscal year; and
<i>STD_RET</i>	= standard deviation of the firm's monthly stock returns from the previous fiscal year.

Analyst Following and the Readability of 10-K Filings

Our first prediction is that analyst following is affected by the level of readability of firms' 10-K filings, as measured by the Fog Index. To control for other factors that can affect analyst following, we estimate the following regression:

$$\begin{aligned} \#ANALYSTS_{i,t} = & \beta_0 + \beta_1 FOG_{i,t} + \beta_2 LOGSIZE_{i,t} + \beta_3 GROWTH_{i,t} + \beta_4 LOGSEGMENTS_{i,t} \\ & + \beta_5 PINST_{i,t} + \beta_6 MFCOUNT_{i,t} + \beta_7 10-K\ NEWS_{i,t} + \beta_8 ADV_{i,t} + \beta_9 R\&D_{i,t} \\ & + \beta_{10} STD\_RET_{i,t} + \varepsilon_{i,t}. \end{aligned} \quad (4)$$

The estimation is performed using ordinary least-squares regression with industry and time fixed effects.

Column 1 of Table 3 (model 1) reports the results of the linear model regression. t-statistics, presented in brackets, are based on standards errors that are robust to heteroscedasticity and are clustered at the firm level. The coefficient on *FOG* is positive and statistically significant, suggesting that analyst following is greater for firms with less readable disclosures. This is consistent with the notion that analysts respond to investors' demand for investment information for firms whose disclosures are more costly to process. The coefficients on the control variables are consistent with prior research. Larger firms are associated with greater analyst following as well as firms with higher growth rates and higher institutional ownership. Consistent with Lang and Lundholm (1996), we find that disclosure practice (*MFCOUNT*) and disclosure informativeness (*10-K NEWS*), are positively related to analyst following. We find that analyst following is negatively associated with the number of business segments.²¹ Similar to Barth et al. (2001), we document that analyst following is greater for firms with higher amounts of advertising and research and development expenses. Consistent with Bhushan (1989), we also find that analyst following is positively associated with firm stock return volatility, suggesting that analysts provide greater support to investors when private information is valuable.

In addition to examining the linear association between *FOG* and analyst following, we also examine the association of analyst following and *FOG* using nonlinear and semi-parametric specifications (second and third columns of Table 3, respectively). In the nonlinear specification (model 2), we add a square term (FOG^2) to Equation (2) to capture the possibility that the association between readability and analyst following differs based on the level of *FOG*. In model 3, we examine a semi-parametric specification by including indicator variables for the various quartiles of *FOG*. The coefficient on FOG^2 (model 2) is negative and significant, indicating that the role of

²¹ This effect becomes negative after controlling for firm size, but is positive if size is omitted. The negative coefficient in the multivariate results is consistent with the findings of Bhushan (1989) and Johnston et al. (2009).

TABLE 3

The Association between Annual Report Readability and Analyst Following

Panel A: OLS Regression

Variable	Model 1	Model 2	Model 3
Intercept	-13.067*** [-16.79]	-20.190*** [-6.29]	-11.625*** [-18.23]
FOG	0.085*** [3.74]	0.798*** [2.58]	
FOG ²		-0.018** [-2.33]	
FOG Q2			0.168** [2.41]
FOG Q3			0.331*** [4.11]
FOG Q4			0.369*** [4.18]
LOGSIZE	2.623*** [62.20]	2.624*** [62.23]	2.623*** [62.22]
GROWTH	0.361*** [3.90]	0.356*** [3.85]	0.356*** [3.85]
LOGSEGMENTS	-0.367*** [-4.68]	-0.368*** [-4.70]	-0.368*** [-4.70]
PINST	0.011*** [5.20]	0.011*** [5.20]	0.011*** [5.22]
MFCOUNT	0.086*** [5.82]	0.086*** [5.82]	0.086*** [5.83]
10-K NEWS	2.391*** [5.46]	2.380*** [5.43]	2.371*** [5.40]
ADV	7.555*** [4.57]	7.564*** [4.57]	7.553*** [4.57]
R&D	1.696*** [5.61]	1.673*** [5.54]	1.662*** [5.50]
STD_RET	0.932*** [2.61]	0.924*** [2.59]	0.931*** [2.61]
Time Fixed Effects	Yes	Yes	Yes
Industry Fixed Effects	Yes	Yes	Yes
n	33704	33704	33704
Adjusted R ²	0.63	0.63	0.63

Panel B: Comparative Interquartile Effects

	Model 1	Model 2	Model 3
FOG	0.146	0.188	0.201
LOGSIZE	6.236	6.238	6.238
GROWTH	0.071	0.070	0.070
LOGSEGMENTS	-0.403	-0.404	-0.405
PINST	0.473	0.473	0.478
MFCOUNT	0.172	0.172	0.172
10-K NEWS	0.081	0.080	0.080
ADV	0.042	0.042	0.042

(continued on next page)

Panel B: Comparative Interquartile Effects

	Model 1	Model 2	Model 3
<i>RD</i>	0.130	0.129	0.128
<i>STD_RET</i>	0.098	0.097	0.174

******, ******* $p < 0.05$ and $p < 0.01$, respectively, two-tailed t-tests.

Panel A reports coefficient estimates and t-statistics (in brackets) from the regression of analyst following on *FOG* and control variables. Analyst following is defined as the number of analysts contained in the first I/B/E/S consensus annual earnings forecast following the 10-K filing. *FOG QX* is an indicator variable set to 1 if the firm's 10-K Fog Index is in quartile *X* of the sample, and 0 otherwise. All other variables are as defined in Table 2. Industry fixed effects are based on the Fama and French 48-industry classification. t-statistics are robust to heteroscedasticity and clustered at the firm level. Panel B reports the estimated effect of an interquartile change in the variable of interest.

readability increases at a declining rate. Also, the coefficients on the second, third, and fourth *FOG* quartiles (model 3) are positive and significant and monotonically decreasing across the quartiles. These results are consistent with analysts making a trade-off between the benefits of covering firms with less readable disclosures and the costs of following them. While these costs include the processing costs of covering firms with less readable disclosures, analysts likely find it more difficult to accurately forecast the earnings of firms with more complex disclosures. These costs can be important, as prior research has found that forecast accuracy has important career effects for analysts (Mikhail et al. 1999; Hong and Kubik 2003). The effects of *FOG* on the properties of analyst earnings forecasts will be examined later.²²

Table 3, Panel B compares the effects of an interquartile change in the independent variables on analyst following based on the marginal effect estimates in Panel A. Consistent with prior literature, size (*LOGSIZE*) has the largest association with analyst following. An interquartile change in the natural logarithm of size is associated with an increase of 6.2 analysts, while that of all other variables is less than 1. The interquartile effect of *FOG* is 0.15 in model 1, 0.19 in model 2, and 0.20 in model 3. While the incremental effect of the complexity of 10-K filings on analyst following is moderate, it is greater than that of most of the other variables included in the regressions. These are the same variables that have been shown in the literature to be related to analyst following. Based on these comparisons, we conclude that disclosure readability is an important determinant of analyst following.

Similar to Lang and Lundholm (1996), we also examine the association between *changes* in *FOG* and *lead changes* in analyst following. If analysts do, in fact, respond to changes in *FOG* by increasing their coverage, then we expect to observe a positive relation between current changes in *FOG* and future changes in analysts following. We investigate this relation by estimating multivariate models using the changes in the independent variables from Table 3. We define the current change in the Fog Index and other independent variables as the difference between their current values and those of the prior fiscal year. We define the lead change in analyst following as the difference between the number of analysts following the firm after the next 10-K filing and the number of analysts following the firm after the current report. Table 4 presents results from these tests. We find that contemporaneous changes in *FOG* are positively related to lead changes in analyst following. We also find that changes in firm size and institutional ownership are positively

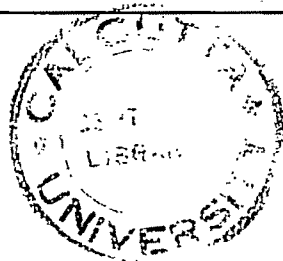
²² In a related study, You and Zhang (2009) examine stock returns following 10-K filings and find that investors underreact to firms with more complex annual filings. This apparent underreaction provides additional motivation for analyst coverage because prior work has found that analysts help improve market efficiency (Barron et al. 2002; Brown and Sivakumar 2003; Gu and Chen 2004).

TABLE 4
The Association between Changes in Annual Report Readability and Lead Changes in Analyst Following

Variable	Model 1	Model 2	Model 3
Intercept	-0.038*** [-3.79]	-0.051*** [-4.68]	-0.073*** [-3.03]
ΔFOG	0.024*** [2.66]	0.023** [2.55]	
$(\Delta FOG)^2$		0.007** [2.58]	
$\Delta FOG Q2$			-0.009 [-0.28]
$\Delta FOG Q3$			0.045 [1.36]
$\Delta FOG Q4$			0.110*** [3.20]
$\Delta LOGSIZE$	0.480*** [21.89]	0.479*** [21.83]	0.480*** [21.91]
$\Delta GROWTH$	0.048 [0.66]	0.045 [0.62]	0.050 [0.68]
$\Delta LOGSEGMENTS$	-0.103** [-2.12]	-0.103** [-2.11]	-0.102** [-2.11]
$\Delta PINST$	0.008*** [5.89]	0.008*** [5.82]	0.008*** [5.84]
$\Delta MFCOUNT$	-0.015 [-1.61]	-0.015 [-1.61]	-0.014 [-1.59]
$\Delta 10-K NEWS$	-0.138 [-0.91]	-0.138 [-0.91]	-0.136 [-0.90]
ΔADV	0.333 [0.50]	0.327 [0.49]	0.323 [0.48]
$\Delta R\&D$	-0.075 [-0.57]	-0.075 [-0.57]	-0.074 [-0.56]
ΔSTD_RET	-0.637*** [-5.18]	-0.631*** [-5.14]	-0.631*** [-5.14]
n	27226	27226	27226
Adjusted R ²	0.03	0.03	0.03

, * p < 0.05 and p < 0.01, respectively, two-tailed t-tests.

This table reports coefficient estimates and t-statistics (in brackets) from the regression of lead changes in analyst following on changes in *FOG* and control variables. Lead change in analyst following is the difference between the number of analysts following the firm after the next 10-K filing and the number of analysts following the firm after the current report. $\Delta FOG QX$ is an indicator variable set to 1 if the current change in a firm's 10-K Fog Index is in quartile X of the sample, and 0 otherwise. All other variables are current period changes of the variables defined in Table 2. t-statistics are robust to heteroscedasticity and clustered at the firm level.



related to lead changes in analyst following and that changes in management earnings guidance, number of business segments, and return volatility are negatively related.²³ In untabulated results, we find that *FOG* has the third-highest effect of the variables employed in our analysis. The results from this test are important because they help to alleviate concerns about endogeneity or prior period information shocks relating to the levels results in Table 3.

Analyst Report Duration

One potential measure of the costs or effort that analysts bear in following firms with less readable disclosures is the amount of time it takes them to issue reports following the 10-K filing. We label this length of time “analyst report duration.” In this section, we examine the analyst report duration at both the analyst and firm level. We expect that analyst report duration will be longer for firms with less readable 10-K filings. As noted previously, to ensure that we capture analyst reports issued in response to the filing of the 10-K (and not in response to other corporate events), we include in this analysis only firm-analyst observations for which the analyst has issued at least one report during the 90 days prior to the 10-K filing and issues a report in the 90 days subsequent to the 10-K filing.²⁴ Table 5, Panel A presents univariate results based on individual analyst reports and the Fog Index of the firms they cover. We classify each firm-specific *FOG* as high (low) if it is greater (less) than the median *FOG* value for all firms in the sample. We find that analysts covering firms with a high *FOG* require 1.45 days longer, on average, to issue their reports. Panels B and C of Table 5 present the average analyst report duration per firm-year and analyst-year, respectively. The value of *FOG* at the firm level is the individual firm 10-K Fog for each specific year. The *FOG* score at the analyst level is the average 10-K Fog score of the firms that the analyst covered in each year.²⁵ The univariate results suggest that the average (median) analyst report duration for firms with *FOG* scores higher than the median is 1.11 (1.50) days (Panel B). At the analyst level, we find that analysts who cover a portfolio of firms with an average *FOG* score higher than the median analyst portfolio average take 2.29 (2.17) days longer on average (median) to issue their first reports. These differences are statistically significant at a 1 percent level and are likely to be economically important given the rapid pace at which markets impound new information.

To further examine the analyst report duration, we estimate ordinary least-squares regressions on the average analyst report durations at both the firm and analyst level. Table 5, Panel D presents the results of these analyses. Similar to the univariate results, we find that firms with higher Fog scores and analysts who cover stocks with higher Fog scores are associated with longer analyst report duration. Specifically, controlling for a variety of other factors that can affect firms’ information environment, the coefficient on *FOG* is 0.345 for the firm-level analysis and 1.138 for the analyst-level analysis, both of which are statistically significant. This suggests that firms and analyst portfolios with one unit higher of *FOG* have analyst response times that are 0.345 and 1.138 days longer. Although it is difficult to assess the economic significance of these results, we note that given the pace with which new analyst information is incorporated in stock prices, even such short delays in producing a report could reduce the information content of the report. The results in Table 5 further indicate that firms with greater institutional ownership, discretionary

²³ While less intuitive, one potential concern is that analyst following might lead *FOG*. For example, management might provide more complex information in response to a greater analyst following. In untabulated tests we find no relation between changes in *FOG* and lag analyst following.

²⁴ As noted previously, we exclude analyst observations made after earnings announcements that follow the 10-K filing and reports made after 90 days after the 10-K. Our results are similar if we use reports within a year after the 10-K filing; however, these observations are likely unrelated to the *FOG* score of the 10-K filing.

²⁵ To be included in our analysis each analyst must cover a minimum of four firms. Similar results are obtained by requiring only a single firm.

TABLE 5

The Association between Annual Report Readability and Analyst Report Duration

Panel A: Individual Analyst Reports

Group	n	Mean	Median
High Fog Index	30,892	19.19	16.00
Low Fog Index	30,884	17.74	15.00
Difference		1.45***	1.00***

Panel B: Average Analyst Report Duration Per Firm

Group	n	Mean	Median
High Fog Index	8,701	18.33	16.75
Low Fog Index	8,701	17.22	15.25
Difference		1.11***	1.50***

Panel C: Average Analyst Report Duration Per Analyst

Group	n	Mean	Median
High Fog Index	2,869	19.91	18.67
Low Fog Index	2,868	17.62	16.50
Difference		2.29***	2.17***

Panel D: Regression Analysis

	Mean Duration at Firm Level	Mean Duration at Analyst Level
Intercept	9.430 ^{††} [7.62]	-8.750 ^{††} [-2.62]
FOG	0.345 ^{††} [5.79]	1.138 ^{††} [6.96]
LOGSIZE	0.034 [0.51]	-0.013 [-0.11]
GROWTH	-0.453 [-1.62]	2.634 ^{††} [3.93]
LOGSEGMENTS	0.405 ^{††} [2.57]	0.844 ^{††} [2.67]
PINST	0.039 ^{††} [9.22]	0.105 ^{††} [11.50]
MFCOUNT	0.170 ^{††} [5.68]	0.019 [0.33]
10-K NEWS	-25.409 ^{††} [-12.33]	-43.743 ^{††} [-8.01]
ADV	10.053 ^{††} [2.52]	30.936 ^{††} [4.06]
R&D	1.925 ^{††} [3.06]	0.544 [0.56]
STD_RET	-6.300 ^{††} [-5.53]	-13.345 ^{††} [-5.86]
n	17868	5737
Adjusted R ²	0.04	0.11

(continued on next page)

*** $p < 0.01$, two-tailed t-tests for differences of means and Wilcoxon for medians.

††† $p < 0.01$, two-tailed t-tests.

Analyst report duration is defined as the length of time between the 10-K filing and the first report for each analyst. Panels A, B, and C report descriptive statistics of the differences in the average analyst report duration at the individual analyst report, firm-year, and analyst-year levels. The firm and analyst levels are computed by averaging variables over the unit of observation. The High and Low groups are based on the median of the sample in each group. Statistical differences of means (medians) are computed based on t-tests (Wilcoxon tests).

Panel D reports coefficient estimates based on ordinary least square regressions at both the firm and analyst levels. All variables are defined as in Table 2. On the analyst level, the variables are averaged over the portfolio of each analyst-year observation. t-statistics (in brackets) are robust to heteroscedasticity and clustered at the firm and analyst levels, respectively.

disclosure, and firm intangibles are associated with longer average analyst response times. In contrast, firms with greater reaction to the release of the 10-K report and greater prior return volatility are associated with shorter analyst report response time.²⁶ Overall, the evidence is consistent with the notion that analysts exert more effort to cover firms with less readable disclosures.

Information Content of Analyst Reports

Our third hypothesis predicts that the informativeness of analyst reports is positively related to the readability of firms' 10-K reports. To test this hypothesis, we estimate the following ordinary least-squares regression with industry and time fixed effects:

$$AI_{i,t} = \beta_0 + \beta_1 FOG_{i,t} + \beta_2 LOGSIZE_{i,t} + \beta_3 GROWTH_{i,t} + \beta_4 LOGSEGMENTS_{i,t} + \beta_5 PINST_{i,t} \\ + \beta_6 MFCOUNT_{i,t} + \beta_7 10-K\ NEWS_{i,t} + \beta_8 ADV_{i,t} + \beta_9 R\&D_{i,t} + \beta_{10} STD\_RET_{i,t} + \varepsilon_{i,t}. \quad (5)$$

Similar to the regression reported in Table 3, we also examine the association of the information content of analyst reports and *FOG* using nonlinear and semi-parametric specifications.

Table 6 reports the regression findings. t-statistics, presented in brackets, are based on standard errors that are robust to heteroscedasticity and clustered at the firm level. Consistent with our hypothesis, the coefficient on *FOG* is positive and significant, suggesting that the informativeness of analyst reports is increasing in the complexity of the 10-K disclosure. This evidence is consistent with the notion that investors find analyst reports for firms with less readable disclosures more useful because of greater processing costs. Analysts' private information searches could also be more valuable in such cases. In model 2, the square term (FOG^2) is negative but insignificant, whereas the coefficients on the individual *FOG* quartiles in model 3 are significant and monotonically increasing. These findings are consistent with our prior results and support the notion that analysts trade off costs and benefits in their coverage of firms with less readable disclosure. The effects of the control variables on the information content of analyst reports are similar in direction to their effects on analyst following. Specifically, the informativeness of analyst reports is increasing in firm size, growth, institutional ownership, discretionary disclosure, firm intangibles, and return volatility. In untabulated results, we find that an interquartile change in *FOG* increases the

²⁶ In untabulated results, we include a control for the number of reports issued by an analyst for the analyst level results. The coefficient on this variable is not statistically significant and does not quantitatively affect the results as presented. We also estimate two proportional hazard models to examine the influence of readability on the time to an analyst's first report by conditioning at the firm and analyst level, similar to O'Brien et al. (2005). We find that, on average, the probability of an analyst making a report at any given time is between 1.7 percent and 3.2 percent lower for each unit change of *FOG*.

TABLE 6

The Association between Annual Report Readability and the Information Content of Analysts' Reports

Variable	Model 1	Model 2	Model 3
Intercept	-0.256*** [-16.90]	-0.302*** [-4.09]	-0.227*** [-20.43]
FOG	0.002*** [3.25]	0.006 [0.86]	
FOG ²		-0.000 [-0.62]	
FOG Q2			0.003** [1.99]
FOG Q3			0.006*** [3.33]
FOG Q4			0.007*** [3.51]
LOGSIZE	0.049*** [65.99]	0.049*** [66.03]	0.049*** [66.02]
GROWTH	0.012*** [5.54]	0.011*** [5.52]	0.011*** [5.50]
LOGSEGMENTS	-0.011*** [-6.47]	-0.011*** [-6.47]	-0.011*** [-6.48]
PINST	0.000*** [10.09]	0.000*** [10.09]	0.000*** [10.12]
MFCOUNT	0.003*** [8.70]	0.003*** [8.70]	0.003*** [8.72]
10-K NEWS	0.049*** [4.34]	0.049*** [4.33]	0.049*** [4.30]
ADV	0.089*** [2.65]	0.090*** [2.65]	0.089*** [2.64]
R&D	0.006 [0.92]	0.005 [0.89]	0.005 [0.83]
STD_RET	0.036*** [4.41]	0.036*** [4.41]	0.036*** [4.42]
Time Fixed Effects	Yes	Yes	Yes
Industry Fixed Effects	Yes	Yes	Yes
n	30716	30716	30716
Adjusted R ²	0.51	0.51	0.51

, * p < 0.05, and p < 0.01, respectively, two-tailed t-tests.

This table reports coefficient estimates and t-statistics (in brackets) from the regression of the information content of analyst reports on *FOG* and control variables. Analysts information content is defined as the proportion of firm stock returns related to analyst forecast revisions to the total firm stock return during the time period between the 10-K filing and the subsequent fiscal year-end. *FOG QX* is an indicator variable set to 1 if the firm's 10-K Fog Index is in quartile *X* of the sample, and 0 otherwise. All other variables are as defined in Table 2. Industry fixed effects are based on the Fama and French 48-industry classification. t-statistics are robust to heteroscedasticity and clustered at the firm level.

information content of analyst reports by about 0.29 percent to 0.46 percent, on average, across the three models. Consistent with our finding for analyst following, we find that the effect of an

interquartile change in disclosure readability (*FOG*) is similar to many of the other variables in our analysis.²⁷

Properties of Analyst Earnings Forecasts

Given our findings regarding the association of disclosure readability with analyst following, analyst report duration, and the information content of analyst reports, a reasonable next step is to examine its association with properties of analyst forecasts. In particular, we focus on the dispersion, accuracy, and the overall and common uncertainty associated with analyst forecasts of annual earnings. As previously explained, we hypothesize that less readable disclosures will be associated with greater forecast dispersion and lower analyst forecast accuracy. We also predict that overall analyst uncertainty will be higher for firms with less readable disclosures, but make no directional prediction regarding analyst common uncertainty.

For each of these hypotheses, we estimate a model of the following form:

$$\begin{aligned} \text{Forecast Property}_{i,t} = & \beta_0 + \beta_1 \text{FOG}_{i,t} + \beta_2 \text{LOGSIZE}_{i,t} + \beta_3 \text{GROWTH}_{i,t} + \beta_4 \text{LOGSEGMENTS}_{i,t} \\ & + \beta_5 \text{PINST}_{i,t} + \beta_6 \text{MFCOUNT}_{i,t} + \beta_7 \text{10-K NEWS}_{i,t} + \beta_8 \text{ADV}_{i,t} \\ & + \beta_9 \text{R\&D}_{i,t} + \beta_{10} \text{STD\_RET}_{i,t} + \varepsilon_{i,t}, \end{aligned} \quad (6)$$

where “Forecast Property” represents the previously defined variable of interest (*DISPERSION*, *ACCURACY*, *UNCERTAINTY_{OVERALL}*, or *UNCERTAINTY_{COMMON}*). Each model is estimated using ordinary least-squares regression with industry and time fixed effects. Table 7 presents the results. t-statistics, presented in brackets, are based on standard errors that are robust to heteroscedasticity and clustered at the firm level.

As reported in the first two columns of Table 7, we find that the coefficient on *FOG* is positive and significant, indicating that less readable disclosures are associated with more dispersed and less accurate analyst earnings forecasts. Consistent with prior research, these results, in conjunction with our findings about the information content of analysts’ reports, suggest that analysts provide information that investors find valuable at the cost of less accurate forecasts (Schipper 1991; Clement and Tse 2003). It is, however, difficult to assess the significance of these benefits because we cannot observe the appropriate benchmark—i.e., the effect on investors of analysts choosing not to follow firms with less readable disclosures.²⁸

We further examine these results by testing the effect of 10-K readability on measures of overall and common analyst forecast uncertainty, as derived by Barron et al. (1998). We report our results in columns three and four of Table 7. Consistent with our prediction, we find that overall analyst uncertainty is increasing in *FOG*, suggesting that there is higher overall uncertainty in the analyst information environment for firms with less readable 10-K filings. Further, we find that the proportion of common analyst uncertainty to overall uncertainty is also increasing in *FOG*, suggesting that as readability decreases, publicly available information, such as the 10-K, becomes

²⁷ As an additional sensitivity test, we follow the suggestion of a reviewer and examine the effect of analysts’ conflicts of interest associated with the Global Research Analyst Settlement to support our inferences. Because analysts with conflicts of interest are likely to place less importance on firm disclosures, we predict that our results for analyst following and report information content should be weaker when analysts’ conflicts of interest are higher. We note, however, that other important events such as Regulation Fair Disclosure occurred around this time and could limit our inferences from this test. In untabulated results, we find that, consistent with our prediction, the association between analyst following and analyst report information content is stronger in the years following the Global Research Analyst Settlement, suggesting greater analyst reliance and attention to 10-K filings in the post period. This test further supports our overall conclusion regarding the association between analyst behavior and 10-K readability.

²⁸ Related to this point, Lee (2010) finds that less readable 10-Q filings are associated with lower information efficiency and greater information asymmetry. She also finds that these negative consequences of lower readability are mitigated by analyst coverage, which provides evidence of the benefits of analysts following firms with less readable reports.

TABLE 7
The Association between Annual Report Readability and Properties of Analysts' Information Environment

Variable	<i>DISPERSION</i>	<i>ACCURACY</i>	<i>UNCERTAINTY_{OVERALL}</i>	<i>UNCERTAINTY_{COMMON}</i>
Intercept	0.0192*** [4.04]	0.2887 [1.35]	-0.0142 [-0.37]	-0.0440 [-0.56]
<i>FOG</i>	0.0003*** [3.96]	0.0041** [2.10]	0.0044*** [2.58]	0.0053** [2.23]
<i>LOGSIZE</i>	-0.0027*** [-22.00]	-0.0315*** [-9.89]	-0.0133*** [-7.19]	-0.0012 [-0.43]
<i>GROWTH</i>	0.0008 [1.08]	0.0077 [0.45]	-0.0005 [-0.04]	0.0638*** [5.69]
<i>LOGSEGMENTS</i>	0.0008*** [3.74]	0.0134** [1.96]	0.0025 [0.54]	0.0081 [1.24]
<i>PINST</i>	-0.0001*** [-8.45]	0.0001 [0.49]	-0.0001 [-0.77]	0.0010*** [5.48]
<i>MFCOUNT</i>	-0.0001*** [-2.92]	0.0010 [1.10]	-0.0001 [-0.11]	0.0034*** [2.79]
<i>10-K NEWS</i>	0.0511*** [7.48]	0.7263*** [5.80]	0.5591*** [4.75]	0.3838*** [4.37]
<i>ADV</i>	0.0266*** [4.13]	0.3901** [2.13]	0.0933 [1.03]	0.0255 [0.18]
<i>R&D</i>	0.0081*** [5.97]	-0.0266 [-0.87]	-0.0036 [-0.14]	-0.1394*** [-4.39]
<i>STD_RET</i>	0.0242*** [9.27]	0.3793*** [6.03]	0.2209*** [4.36]	0.3683*** [6.72]
Time Fixed Effects	Yes	Yes	Yes	Yes
Industry Fixed Effects	Yes	Yes	Yes	Yes
n	26078	29055	17241	17222
Adjusted R ²	0.20	0.05	0.04	0.06

, * p < 0.05, and p < 0.01, respectively, two-tailed t-tests.

This table reports coefficient estimates and t-statistics (in brackets) from regressions of analyst earnings forecast dispersion, accuracy, and uncertainty on *FOG* and control variables. Analyst earnings forecast dispersion is defined as the standard deviation of the individual analyst earnings forecasts in the most recent consensus earnings forecast following the annual report scaled by price. Analyst earnings forecast accuracy is defined as the squared value of the difference between the I/B/E/S actual reported earnings and the most recent analyst consensus earnings forecast following the annual report filing scaled by price. Analyst overall and common uncertainty are calculated following the equations derived by Barron et al. (1998). All other variables are as defined in Table 2. Industry fixed effects are based on the Fama and French 48-industry classification. t-statistics are robust to heteroscedasticity and clustered at the firm level.

more important to analysts relative to private information. This suggests that analysts have difficulty producing private information for firms with less readable reports. Overall, the evidence from the four regressions in Table 7 suggests that analyst earnings forecasts are affected by 10-K readability, supporting the notion that analysts use the information contained in the 10-K filings.

The results for the control variables are largely consistent with the notion that analysts perform better for firms with better information environments and less complex operations. We find that firms with better information environments, as measured by size and institutional ownership, are associated with greater accuracy, lower dispersion, and lower uncertainty. Institutional ownership is also positively associated with analyst common uncertainty, consistent with these inves-

tors improving the importance of public information. We find that firms with greater uncertainty and complexity, as measured by the volatility of returns, the number of business segments, and the extent of intangible investment, are generally associated with lower forecast accuracy, greater dispersion, and greater overall uncertainty. However, their relation to analyst common uncertainty is less clear. While *R&D* investment is negatively related to common uncertainty (because less public information on these investments is available), return volatility is positively related. These findings suggest that private information is less important relative to common information for firms with high return volatility. While the measures of firm disclosure, *MFCOUNT* and *10-K NEWS*, are both positively associated with analyst common uncertainty (i.e., public disclosures increase the importance of common information), they differ with respect to analyst forecast dispersion. *MFCOUNT* is associated with lower forecast dispersion, consistent with managers providing guidance to reduce analyst disagreement, while *10-K NEWS* is associated with greater dispersion and lower accuracy.

V. SUMMARY AND CONCLUSION

From the passage of the Securities Acts in 1933 and 1934 to the present, regulators, legal scholars, and various other parties have weighed in on the debate about the complexity of financial communication to external users. The SEC has gone so far as to require that the prospectuses of all registered public offerings meet the requirements of the Plain English Rules, and has suggested that similar procedures should be applied to other mandatory filings. Former SEC Chairman Christopher Cox has even suggested the use of readability models such as the Fog Index to measure the complexity of financial communication (Cox 2007). However, other users are concerned about the effects that these types of actions could have on the disclosure regime. These parties argue that, at a minimum, these actions are a waste of time and effort, and could lead to a reduction in disclosure as firms are forced to simplify their public communication (e.g., Kripke 1970, 1973; Firtel 1999). This study sheds insight relevant to this debate by examining the effect of readability on one important financial information intermediary—sell-side financial analysts.

We find evidence consistent with the notion that, because less readable firm communication is more costly to process and interpret, investors demand greater amounts of analyst services for firms with less readable communication. We find that analyst following is greater for firms with higher levels of syntactic complexity as measured by the Fog Index. We also find that analysts who cover firms with less readable communication take longer time on average to issue reports in response to 10-K filings. We interpret this evidence as analysts exerting greater effort to cover these firms. In addition, we find that analyst reports of firms with less readable 10-K reports are more informative to investors, but that the earnings forecasts of such firms have greater analyst dispersion or disagreement, are less accurate, and are associated with greater levels of uncertainty.

Our results suggest that analyst behavior is related to the readability of firms' communication. While prior studies have found that analysts are affected by the complexity of individual financial items, we provide evidence that the *overall* linguistic complexity of firms' communication incrementally influences analyst behavior over and above the effects of the content of the document (e.g., taxes, interest rates). Finally, our results that analysts provide greater amounts of information to investors for firms with less readable communication and that investors consider this information informative are relevant to the SEC's debate about the intended audience of financial information and the SEC's concerns on the accessibility of these reports. While the SEC moves to reduce the complexity of firms' communication, further research is needed to examine the explicit costs and benefits of such actions. As a final caveat, while the collection of the evidence from our various analyses of analysts' behavior conducted in this study is consistent with the prediction that 10-K readability influences analyst behavior, our individual findings still are subject to the limitation that the documented relations reflect associations and may not be fully causal.

REFERENCES

- Armstrong, J. S. 1980. Unintelligible management research and academic prestige. *Interfaces* 10 (2): 80–86.
- Ball, R. 1992. The earnings price anomaly. *Journal of Accounting and Economics* 15 (2–3): 319–345.
- Barron, O. E., D. Byard, O. Kim, and E. J. Riedl. 2002. High-technology intangibles and analysts' forecasts. *Journal of Accounting Research* 40 (2): 289–312.
- , ———, O. Kim, and D. E. Stevens. 1998. Using analysts' forecasts to measure properties of analysts' information environment. *The Accounting Review* 73 (4): 421–433.
- Barth, M. E., R. Kasznik, and M. F. McNichols. 2001. Analyst coverage and intangible assets. *Journal of Accounting Research* 39 (1): 1–34.
- Bhushan, R. 1989. Firm characteristics and analyst following. *Journal of Accounting and Economics* 11 (2–3): 255–274.
- Biddle, G. C., G. Hilary, and R. S. Verdi. 2009. How does financial reporting quality relate to investment efficiency? *Journal of Accounting and Economics* 48 (2–3): 112–131.
- Bloomfield, R. J. 2002. The “incomplete revelation hypothesis” and financial reporting. *Accounting Horizons* 16 (3): 233–243.
- Botosan, C. A., and M. S. Harris. 2000. Motivations for a change in disclosure frequency and its consequences: An examination of voluntary quarterly segment disclosures. *Journal of Accounting Research* 38 (2): 329–353.
- Bradshaw, M. T., G. S. Miller, and G. Serafeim. 2008. Accounting method heterogeneity and analysts' forecasts. Working paper, Harvard University and University of Michigan.
- Brennan, M. J., and P. J. Hughes. 1991. Stock prices and the supply of information. *The Journal of Finance* 46 (5): 1665–1691.
- , and A. Subrahmanyam. 1995. Investment analysis and price formation in securities markets. *Journal of Financial Economics* 38 (3): 361–381.
- Brown, L. D., and S. Sivakumar. 2003. Comparing the value relevance of two operating income measures. *Review of Accounting Studies* 8 (4): 561–572.
- Callen, J. L., M. Khan, and H. Lu. 2009. Accounting quality, stock price delay and future stock returns. Working paper, University of Toronto and Massachusetts Institute of Technology.
- Charrow, R. P., and V. R. Charrow. 1979. Making legal language understandable—Psycholinguistic study of jury instructions. *Columbia Law Review* 79 (7): 1306–1374.
- Chomsky, N. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
- Clement, M. B., and S. Y. Tse. 2003. Do investors respond to analysts' forecast revisions as if forecast accuracy is all that matters? *The Accounting Review* 78 (1): 227–249.
- Core, J. E. 2001. A review of the empirical disclosure literature: Discussion. *Journal of Accounting and Economics* 31 (1–2): 441–456.
- Cotter, J., I. Tuna, and P. D. Wysocki. 2006. Expectations management and beatable targets: How do analysts react to explicit earnings guidance? *Contemporary Accounting Research* 23 (3): 593–628.
- Courtis, J. K. 1995. Readability of annual reports: Western versus Asian evidence. *Accounting, Auditing & Accountability Journal* 8 (2): 4–17.
- Cox, C. 2007. Closing remarks to the Second Annual Corporate Governance Summit. Delivered at the USC Marshall School of Business, Los Angeles, CA, March 23. Available at <http://www.sec.gov/news/speech/2007/spch032307cc.htm>.
- Duru, A., and D. M. Reeb. 2002. International diversification and analysts' forecast accuracy and bias. *The Accounting Review* 77 (2): 415–433.
- Firtel, K. B. 1999. Plain English: A reappraisal of the intended audience of disclosure under the Securities Act of 1933. *Southern California Law Review* 72 (2–3): 851–897.
- Frankel, R., S. P. Kothari, and J. Weber. 2006. Determinants of the informativeness of analyst research. *Journal of Accounting and Economics* 41 (1–2): 29–54.
- Gazmararian, J. A., D. W. Baker, and M. V. Williams. 1999. Health literacy among Medicare enrollees in a managed care organization. *Journal of the American Medical Association* 281 (6): 545–551.
- Gibson, E. 1998. Linguistic complexity: Locality of syntactic dependencies. *Cognition* 68 (1): 1–76.
- Gu, Z. Y., and T. Chen. 2004. Analysts' treatment of nonrecurring items in street earnings. *Journal of Accounting and Economics* 38 (1–3): 129–170.

- , and W. M. Wang. 2005. Intangible assets, information complexity, and analysts' earnings forecasts. *Journal of Business Finance & Accounting* 32 (9–10): 1673–1702.
- Gunning, R. 1969. The Fog Index after twenty years. *Journal of Business Communication* 6 (2): 3–13.
- Healy, P., 1977. Can you understand the footnotes to financial statements? *Accountants Journal* July: 219–222.
- Healy, P. M., A. P. Hutton, and K. G. Palepu. 1999. Stock performance and intermediation changes surrounding sustained increases in disclosure. *Contemporary Accounting Research* 16 (3): 485–520.
- Hirst, D. E., and P. E. Hopkins. 1998. Comprehensive income reporting and analysts' valuation judgments. *Journal of Accounting Research* 36: 47–75.
- , ———, and J. M. Wahlen. 2004. Fair values, income measurement, and bank analysts' risk and valuation judgments. *The Accounting Review* 79 (2): 453–472.
- Hong, H., and J. D. Kubik. 2003. Analyzing the analysts: Career concerns and biased earnings forecasts. *The Journal of Finance* 58 (1): 313–351.
- Hopkins, P. E., R. W. Houston, and M. F. Peters. 2000. Purchase, pooling, and equity analysts' valuation judgments. *The Accounting Review* 75 (3): 257–281.
- Indjejikian, R. J. 1991. The impact of costly information interpretation on firm disclosure decisions. *Journal of Accounting Research* 29 (2): 277–301.
- Johnston, R., S. Markov, and S. Ramnath. 2009. Sell-side debt analysts. *Journal of Accounting and Economics* 47: 91–107.
- Jones, M. J., and P. A. Shoemaker. 1994. Accounting narratives: A review of empirical studies of content and readability. *Journal of Accounting Literature* 13: 142–184.
- Kimble, J. 1994. Answering the critics of plain language. *The Scribes Journal of Legal Writing* 5: 51–85.
- Koo, M. M., I. Krass, and P. Aslani. 2003. Factors influencing consumer use of written drug information. *The Annals of Pharmacotherapy* 37 (2): 259–267.
- Kripke, H. 1970. The SEC, the accountants, some myths and some realities. *New York University Law Review* 45 (6): 1151–1205.
- . 1973. The myth of the informed layman. *Business Law* 28: 631–638.
- Lang, M. H., and R. J. Lundholm. 1996. Corporate disclosure policy and analyst behavior. *The Accounting Review* 71 (4): 467–492.
- , K. V. Lins, and D. P. Miller. 2004. Concentrated control, analyst following, and valuation: Do analysts matter most when investors are protected least? *Journal of Accounting Research* 42 (3): 589–623.
- Lee, S. Y. D., J. A. Gazmararian, and A. M. Arozullah. 2006. Health literacy and social support among elderly Medicare enrollees in a managed care plan. *Journal of Applied Gerontology* 25 (4): 324–337.
- Lee, Y. 2010. The effects of quarterly reported readability on information efficiency of stock prices. Working paper, National Taiwan University.
- Leuz, C., and C. M. Schrand. 2009. Disclosure and the cost of capital: Evidence from firms' response to the Enron shock. Working paper, University of Chicago and University of Pennsylvania.
- Li, F. 2008. Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and Economics* 45 (2–3): 221–247.
- Li, Y., P. R. Rau, and J. Xu. 2009. All-star analyst coverage and corporate transparency. Working paper, Purdue University.
- Loughran, T., and B. McDonald. 2010. Measuring readability in financial text. Working paper, University of Notre Dame.
- Mikhail, M. B., B. R. Walther, and R. H. Willis. 1999. Does forecast accuracy matter to security analysts? *The Accounting Review* 74 (2): 185–200.
- Miller, P. B. 2010. The effects of reporting complexity on small and large investor trading. *The Accounting Review* 85 (6): 2107–2143.
- Nagar, V., D. Nanda, and P. Wysocki. 2003. Discretionary disclosure and stock-based incentives. *Journal of Accounting and Economics* 34 (1–3): 283–309.
- O'Brien, P. C., M. F. McNichols, and H. W. Lin. 2005. Analyst impartiality and investment banking relationships. *Journal of Accounting Research* 43 (4): 623–650.
- , and R. Bhushan. 1990. Analyst following and institutional ownership. *Journal of Accounting Research* 28: 55–76.

- Ott, B. B., and T. L. Hardie. 1997. Readability of advance directive documents. *Image—the Journal of Nursing Scholarship* 29 (1): 53–57.
- Phillips, T. J., C. M. Daily, and M. S. Luehlfigg. 2007. A note on the readability of professional materials for management accountants. *Advances in Management Accounting* 16: 311–318.
- Plumlee, M. A. 2003. The effect of information complexity on analysts' use of that information. *The Accounting Review* 78 (1): 275–296.
- Previts, G. J., R. J. Bricker, T. R. Robinson, and S. J. Young. 1994. A content analysis of sell-side financial analysts company reports. *Accounting Horizons* 8 (2): 55–70.
- Rogers, R. K., and J. Grant. 1997. Content analysis of information cited in reports of sell-side financial analysts. *Journal of Financial Statement Analysis* 3 (1): 17–30.
- Schipper, K. 1991. Analysts' forecasts. *Accounting Horizons* 5 (4): 105–121.
- Schroeder, M. 2002. SEC proposes rules to improve disclosure by public companies. *Wall Street Journal* (May 1): C5.
- Schroeder, N. and C. Gibson. 1990. Readability of management's discussion and analysis. *Accounting Horizons* 4 (4): 78–87.
- Securities and Exchange Commission (SEC). 1998. *A Plain English Handbook: How to Create Clear SEC Disclosure Documents*. Washington, D.C.: U.S. Securities and Exchange Commission.
- Shuptrine, F. K., and E. M. Moore. 1980. Even after the Magnuson-Moss Act of 1975, warranties are not easy to understand. *The Journal of Consumer Affairs* 14 (2): 394–404.
- Smith, J. E., and N. P. Smith. 1971. Readability: A measure of the performance of the communication function of financial reporting. *The Accounting Review* 46 (3): 552–561.
- Stone, C. A., E. R. Silliman, B. J. Ehren, and K. Apel. 2005. *Handbook of Language and Literacy: Development and Disorders*. New York, NY: Guilford Press.
- Thompson, C. K., and Y. Faruqi-Shah. 2002. Models of sentence production. In *Handbook of Adult Language Disorders: Integrating Cognitive Neuropsychology, Neurology, and Rehabilitation*, edited by A. E. Hillis, 311–330. Philadelphia, PA: Psychology Press, Taylor and Francis.
- You, H., and X. Zhang. 2009. Financial reporting complexity and investor underreaction to 10-K information. *Review of Accounting Studies* 14 (4): 559–586.

BOOK REVIEWS

Stephen A. Zeff, Editor

Editor's note: Two copies of books for review should be sent to the Book Review Editor: Stephen A. Zeff, Rice University, Jesse H. Jones Graduate School of Business, 6100 Main St., Houston, TX 77005. The policy of *The Accounting Review* is to publish only those reviews solicited by the Book Review Editor. Unsolicited reviews will not be accepted.

JOAN LUFT and MICHAEL D. SHIELDS, *Psychology Models of Management Accounting, Foundations and Trends*[®] in Accounting (Hanover, MA: now Publications, Inc., 2009, ISBN 978-1-60198-346-6, Vol. 4, No. 3–4, pp. 151).

This monograph is the newest in the series entitled “Foundations and Trends[®] in Accounting.” The objective of the series is to provide frameworks for understanding and organizing the vast array of existing literature in various fields of accounting research. This issue is co-authored by Joan Luft and Mike Shields, two thought leaders in the field of psychology-based research in management accounting. The main objective of the monograph is to devise a framework that links what we know about the subjective decisions made by management accountants with the theories in psychology that have proven useful to our understanding of how these decisions are made. These authors have previously developed a comprehensive study of the causes and consequences of management accounting based on an analysis of 275 papers across six leading journals (see Luft and Shields 2003). The current monograph provides a framework for understanding a much more specific and more focused piece of the management accounting literature, that is, the psychological processes underlying the subjective decisions required of management accountants in practice.

Given the vast array of possible pieces of the literature to review, the reader might ask, “Why is this book needed now?” Interest in the field of behavioral decision research in accounting has spiked due at least in part to the awarding of the Nobel Prize in Economic Sciences in 2002 to Daniel Kahneman and Amos Tversky for their work on the psychology of judgment and decision-making and its application to the field of behavioral economics. Although much psychology-based research in management accounting took place before 2002, the recognition of behavioral economics as an important economic discipline seems to have moved behavioral decision theory in management accounting to the forefront. The literature reviewed in this monograph reflects the best of what has been published in this area since then.

This monograph should prove useful to management accounting scholars and doctoral students interested in behavioral decision theory and its applications in management accounting. Doctoral students would benefit from the authors' framework for thinking about subjective decision making in management accounting. In addition, the organized framework allows scholars to more easily identify gaps in the literature that they may wish to fill.

Do not be confused, as this is not a one-to-one mapping of decision-making categories to psychological theories; in fact, some psychological theories are discussed as underpinnings of several of the decision categories identified by the authors. Such confusion may be driven in part by the title—*Psychology Models of Management Accounting*. The authors do discuss models, but I would argue they are management accounting models of subjective decision making. I might re-title the monograph “How Psychological Theory Informs our Understanding of Subjective Decision Making in Management Accounting.”

The monograph is structured in seven chapters plus an appendix. The authors have a clear and concise writing style reflective of many years of academic writing experience. Chapter 1 provides an introduction to the monograph's purpose and objectives as well as an overview of the authors' organizing framework for subjective decision making in management accounting. The organizing framework consists of four categories of subjective

decision making in management accounting and two emerging themes. The authors argue that there is robust evidence consistent with psychological theory underlying the four categories, while there is some interesting, but perhaps less convincing (or at least less well developed), evidence to support the two emerging themes.

Chapter 2 is meant to describe the methodology used by the authors to develop their categorization scheme. The most important information which the reader needs to understand, namely the methodology, is unfortunately relegated to the appendix and the last paragraph of Chapter 7. The authors instead devote much of Chapter 2 to an argument in favor of cognitive task analysis (i.e., a description and analysis of the knowledge and cognitive activities required for high performance on specific tasks) in management accounting. The authors argue that we have not made a point as a discipline to perform detailed cognitive task analyses of the type reported in the audit literature. Consequently, any framework for understanding how psychology affects management accounting decisions must be organized in an alternate form. Although this issue seems somewhat beside the point in the context of the chapter material, the authors' argument concerning the importance of cognitive task analysis (or lack thereof) in management accounting is well taken. The field would benefit from a knowledge of the detailed cognitive processes involved in subjective decision making in management accounting.

At this point, the reader should be aware of the importance of the appendix in order to obtain a good understanding of the authors' methodology. The appendix walks the reader through two examples, one decision-facilitating and one decision-influencing, of the use of management accounting information in subjective decision making. The decision-facilitating example is based on the "grand model" of product capacity planning and pricing developed by Balakrishnan and Sivaramakrishnan (2002). The decision-influencing example references the work of Feltham and Xie (1994), who model the weights placed on performance measures in a multi-measure performance evaluation scheme. In both cases, the authors point out the many individual (subjective) decisions which managers must make in the course of capacity planning and performance evaluation. The implication is that the authors grouped the various individual decisions made in the process of capacity planning and performance evaluation into the categories listed in Chapter 2.

While the examples provided in the appendix are somewhat informative, the link between the examples and the categories derived from them is rather indirect. The authors are not particularly forthcoming about how the appendix maps onto the categories; consequently, the reader must infer the mapping. While this would not be too difficult for experienced management accounting scholars, doctoral students or non-management accounting researchers would have a more difficult time understanding how the lists provided in the appendix distill down to the four categories identified by the authors in developing their framework, potentially limiting its usefulness.

Chapters 3–6 are devoted to discussions of each of the four subjective decision-making categories. For each category, the authors describe the important elements of the decision process and identify relatively recent studies that examine the type of decision under consideration. The authors do a nice job of summarizing each paper and its potential contribution as well as how later papers build on earlier papers to fill the gaps in our understanding. Chapter 3 examines subjective valuation of management accounting-related monetary payoffs, linking these mainly to cognitive psychology, including framing and prospect theory. Chapter 4 examines subjective valuation of non-monetary payoffs, including fairness, honesty, trust, and reciprocity, mainly supported by theories of social and relational exchange drawn from social psychology. Chapter 5 examines the relatively simple subjective decision models used to perform many management accounting-related tasks (e.g., subjective linear programming, mental models, non-compensatory decision models, effects of opportunity, and sunk costs), while Chapter 6 examines models of parameter and variable acquisition and subjective estimation (e.g., subjective estimation of model parameters using accounting information, effect of accounting knowledge, and experience on this subjective estimation). Each of Chapters 3–6 also includes an exhibit indicating the link between causes (i.e., management accounting independent variables) and effects (i.e., management accounting dependent variables) through particular psychological theories. In addition, the exhibits indicate which of the management accounting studies reviewed in the chapter address the links identified in the exhibit, thus providing nice visual summaries of the main points raised in each of these chapters.

After reading these chapters, I wondered how the particular illustrative articles were selected by the authors for inclusion in the monograph. Most of the illustrative articles were published since 2000, and all were published in top-tier accounting journals. The list of articles included in the monograph is by no means exhaustive, nor do I believe the authors meant it to be. The authors justify their focus on recent work because it reflects the "growth and dynamism of psychology-based research" on management accounting (p. 124). This seems reasonable. In addition, I suspect the authors selected the papers based on their extensive collective experience teaching and researching subjective decision making in management accounting.

Finally, Chapter 7 discusses the two identified emerging themes to motivate future work. The emerging themes are (1) that two or three distinctly different subjective models account for most management accounting decisions, and (2) the important differences between intuitive or "gut-feel" decision making (or, as psychologists have labeled them, "System 1 cognitions") and deliberative or consciously controlled decision making (i.e., "System 2 cognitions") and implications of these for management accounting. According to the authors, these areas are ripe for future research. In addition, the authors identify several important, but relatively unexplored, research questions that could benefit from the application of insights from psychology, including:

- Why not study larger, more complex decision models that better reflect real-world decision making?
- What drives selective information acquisition, given that modern information systems can provide so many different pieces of information?
- Why don't we see more non-laboratory studies similar to Ittner et al. (2003) to examine management accounting subjective decision making in real-world settings?

In conclusion, I believe this book provides a good primer for scholars new to management accounting behavioral research on subjective decision making. For scholars already working in the field, the authors provide an organizing framework for thinking about the vast array of work in this area. This allows us to figure out where our research fits within the current literature and to pinpoint areas where future work is needed to move the literature ahead.

For many years, we have talked about integrating economics and psychology-based research in management accounting. Because of the increased level of interest in behavioral economics and behavioral finance, many economics-based researchers are becoming more interested in identifying behavioral theories that would allow them to gain additional insight into their research questions of interest. My hope is that this increased interest leads to fruitful partnerships between economics- and psychology-based researchers in management accounting. In my opinion, this monograph provides a much-needed tool to inform economics-based researchers of the "state of the art" in subjective decision-making research in management accounting. Consequently, this monograph should be helpful in further encouraging such partnerships.

REFERENCES

- Balakrishnan, R., and K. Sivaramakrishnan. 2002. A critical overview of the use of full-cost data for planning and pricing. *Journal of Management Accounting Research* 14: 3–31.
- Feltham, G., and J. Xie. 1994. Performance measure congruity and diversity in multi-task principal/agent relations. *The Accounting Review* 69 (3): 429–453.
- Ittner, C. D., D. F. Larcker, and M. W. Meyer. 2003. Subjectivity and the weighting of performance measures: Evidence from a balanced scorecard. *The Accounting Review* 78 (3): 725–758.
- Luft, J., and M. D. Shields. 2003. Mapping management accounting: Graphics and guidelines for theory-consistent empirical research. *Accounting, Organizations and Society* 28: 169–249.

THERESA LIBBY
Professor of Accounting
University of Waterloo

CHRIS POULLAOS and SUKI SIAN (editors), *Accountancy and Empire: The British Legacy of Professional Organization* (New York, NY: Routledge, 2010, ISBN 978-0-415-45771-2, pp. xv, 263).

This book discusses the historical development of the accounting profession in different parts of the British Empire. Its objective is to address a gap in the literature "by tracing the rise of professional accountancy in selected colonial outposts and examining the role of the institutions regulating it, within the context of the empire experience" (p. 2). It includes contributions of expert academics, some of them having published extensively on the development of accounting professions in different contexts, others focusing their research on the region that they discuss in their contribution to this work. Some of the chapters build upon previously published work, whereas others try to explore new aspects of the history of accountancy. Some of the countries that are discussed in this book are also found in Poullaos (2009), which can serve as a more general introduction to the history of accountancy as a profession.

Although the different chapters clearly focus on the same type of issues, their development is rather variable. This is partly a consequence of the different historical methodologies on which they are based. In Chapter 2 (on South Africa, Australia, and Canada), Chris Poullaos has relied on an extensive set of primary sources from, e.g., the ICAEW archives, the National Archives, and the Scottish Records Office. It is the only contribution that relies heavily on unpublished archival sources. Suki Sian (Chapter 10, on Kenya) and Shraddha Verma (Chapter 9, on India) use a mixture of primary and secondary sources. Chibuike U. Uche (Chapter 4, on Nigeria) refers also to

email correspondence with representatives of professional bodies. The contribution of Devi S. Susela (Chapter 5, on Malaysia) is based on 65 interviews, whereas P. W. Senarath Yapa (Chapter 6, on Sri Lanka) uses information from nine unstructured interviews. The other contributions rely mainly on secondary sources.

The book covers a range of countries, including the self-governing (white) settler colonies, South Africa, Australia, and Canada, and non-settler colonies such as Nigeria, Malaysia, Sri Lanka, Jamaica, Trinidad and Tobago, India, and Kenya. Consequently, it covers most regions of the British Empire, the only notable exception being what could be considered the Arab part of the Empire. Although the current selection of former colonies provides contrasting examples of the uptake of the British models, it would be interesting to study whether the situation in protectorates such as British Somaliland, Yemen, and Oman was any different from that in the other parts of the Empire.

The subtitle of the book, *"The British Legacy of Professional Organization,"* is highly relevant, as it clarifies that the focus of the contributions is on the organization of accountancy in occupational groupings, not on the development of accounting and auditing practices. Only in some chapters is attention given to the development of individual audit firms, most notably in the chapter on Trinidad and Tobago. All chapters provide some background on the countries involved. For most of them, the information provided is sufficient to understand the issues at stake, but in some chapters more circumstantial information would make the reading easier. Most chapters center strongly on the accounting profession, and one wonders whether similar discussions and developments could be found in other professions.

In most of the chapters, the key issue is the development of a local accounting profession and the relation between British accountancy bodies and overseas associations. An overall impression after reading the book is that this was to a large extent a one-way relation, with, in a first stage, British accountants being "exported" to the different parts of the Empire and, in a later phase, locally trained accountants who organized themselves, very often following British models and with direct or indirect interventions by the U.K. bodies. Even the competition that existed between the U.K. bodies in their home market had a clear impact on developments overseas. This is also apparent in the concluding section, where Poullaos and Sian address how the organization of accountants in different parts of the empire has been affected by the empire experience. This section also focuses on the relation between British capital and the development of colonial professional organization. It explains how South Africa, Australia, and Canada were "decolonized" earlier than the other colonies in terms of formal control over accountancy. In many of the non-settler colonies, local legislation confirmed the continued impact of British reporting, auditing, and professional requirements. It led to professional exclusion processes both before and after political independence. In these countries, local accountants typically could organize themselves only after decolonization. The closing section also discusses how the empire experience had an impact on the organization of accountancy in Britain. However, as the different chapters show, this impact seems limited: contrary to the many British expatriate accountants who went to, for example, Kenya, the opposite flow of trainees was very small. An interesting extension of the current research might focus on analyzing quantitative data on, e.g., the number of non-British citizens who were members of the Institute of Chartered Accountants in England and Wales and other U.K. bodies (see also Briston and Kedsle 1997). Some chapters give the impression that they were very rare (e.g., the chapter on Nigeria), but others indicate that there were "many locals" who went to the U.K. for professional qualification (e.g., the chapter on Malaysia).

Apart from the introductory and concluding chapters, each chapter focuses on a group of related colonies or dominions or on an individual country. Poullaos discusses the evolution of the imperial accountancy arena during the 1920s, focusing on South Africa, Australia, and Canada. He extends the earlier work of Chua and Poullaos (2002) on the interactions between British accountants and their associations on the one hand, and dominion accountants on the other hand. Issues at stake include the use of the "chartered accountant" designation, access by British accountants to dominion markets and of dominion accountants to British designations, and participation in British examinations without having to travel to Britain. The chapter provides a quite detailed account of initiatives, events, and negotiations and demonstrates how, by 1930, South African, Canadian, and Australian accountants in different ways had strengthened their position *vis-à-vis* the British profession.

Canada is also the subject of the second chapter, but Alan J. Richardson draws attention to the fact that the British Empire was not the only imperial power that had an impact on the development of accountancy in Canada. He argues that the Canadian profession was shaped by French, British, and American institutional legacies and that the impact of each of these changed over time. These competing influences still exist today, and their inconsistencies allowed Canada to go its own way. Changing professional designations (chartered accountants, certified public accountants ...) exemplify different paths in the development of the profession. Although, overall, the American influence increased, this influence seems to have been more on the standard-setting processes than on the development of professional structures. The chapter ends with a number of examples in which Richardson shows Canada's position as a "middle power" with a distinct point of view on many issues.

The chapter on Jamaica makes it clear that the power struggle between the U.K. institutes was exported to the colonies. This is most obvious in the opposition between the Association of Certified Corporate Accountants (as

the ACCA was then known) and the other bodies. Different professional bodies coexisted, some linked with U.K. bodies, others being led by indigenous accountants. Again, professional designations were important, with conflicts between "chartered" and "certified" accountants that were often related to attempts to become more independent of the British accounting bodies. Specific issues that are discussed in the chapter are the attempt to create an accounting degree at the University of the West Indies that would serve as an entry qualification for the profession, and the entry of the American Institute of Certified Public Accountants. The ACCA played a similar role in Trinidad and Tobago (T&T). Marcia Annisette's contribution focuses on the evolution of the practice link between Britain and T&T. She starts from the dominance of the large multinational audit firms in the 1990s and traces their history in T&T. There were two different phases in the internationalization of accounting practice, the first one linked with colonialism and British accountancy, the second one in the 1960s linked with U.S.-based multinational practices. This increasing impact of international networks should be seen in a context of growing international standardization of accountancy practices.

There are two contributions on African countries, in which Uche discusses Nigeria and Sian treats Kenya. Uche observes that, around the mid 20th century, Nigeria's export trade was dominated by European firms and that expatriate (U.K.) accountants were the only ones present in the country. This colonial influence still continues and impedes training second-tier accountants who could provide services for smaller local business. Only with the coming of independence did Nigerians become qualified as accountants. It is interesting to note that, again, the ACCA played a major role in this, probably due to the structure of its training programs. Uche reports in detail the continuing impact of the ACCA and its interactions with the formation of local professional bodies. In a similar way, Kenya could not break radically with its colonial past following independence in 1963. Its situation was quite comparable to that of Nigeria, a difference being a stronger economic influence of Asians in the former. Sian suggests that the creation of the Association of Accountants in East Africa, acting as an interest group for British accountants, might have been an attempt to close off the market for non-whites. After independence, movements toward an indigenous profession remained influenced by British concepts of professionalism and British imperial history, although eventually an autonomous profession emerged.

Asian countries included are Malaysia, Sri Lanka, and India. Also in these countries, the continuing impact of British accountants and resistance against this influence had a profound impact on the development of accountancy. The development of the profession in Malaysia is characterized by the complex relation between the Malaysian Association of Certified Public Accountants, the Malaysian Institute of Accountants (a government body), and British professional bodies. This relation seems quite interesting, and it shows some specific characteristics. One of these is that Malaysia was not allowed to use the CA designation, a point that is not fully explained in the chapter. The situation in Sri Lanka seems at first rather comparable, but then again different. Competition there was mainly between the Institute of Chartered Accountants of Sri Lanka (ICASL, having members with British qualifications) and the Registered Accountants (local graduates, considered to be the lower tier of the profession). The ICASL managed to reserve the public practice market to a limited number of elite accountants, a closure that it could maintain until 1977. The Indian professional model was also based on the British model. However, interventions by the Indian government were significant, and they resulted in changes in the regulation of the profession. The chapter provides an interesting summary of this interaction between state and profession. Also here, the use of the "chartered" designation was an issue and seems linked with the social stratification within the profession.

The issue of the designation of accounting professionals can perhaps be seen as a red thread throughout the book. Many debates on the relation between British accountants, British accountancy bodies, and local professionals seem to translate into claims on the "Chartered Accountant" designation that apparently served as a quality label (see also Parker 2005).

In conclusion, this book provides an interesting overview of how professional structures and concepts that were introduced under the rule of the British Empire continued to have an impact after the independence of the countries involved. For scholars who want to understand the current status and organization of the accounting profession in these countries, this book is essential reading. Generally speaking, the way these professions are organized is shaped both by a continuing colonial legacy and by the way local governments and professionals have interacted with these colonial remains in order to create an autonomous profession of indigenous accountants.

REFERENCES

- Briston, R. J., and M. J. M. Kedslie. 1997. The internationalization of British professional accounting: The role of the examination exporting bodies. *Accounting Business & Financial History* 7 (2): 175–194.
- Chua, W. F., and C. Poullaos. 2002. The Empire strikes back? An exploration of centre-periphery interaction between the ICAEW and accounting associations in the self-governing colonies of Australia, Canada and South Africa, 1880–1907. *Accounting, Organizations and Society* 27 (4/5): 409–445.

EDITORIAL POLICY AND STYLE INFORMATION

EDITORIAL POLICY

According to the policies set by the Publications Committee (which were endorsed by the Executive Committee and were published in the *Accounting Education News*, June 1987), *The Accounting Review* “should be viewed as the premier journal for publishing articles reporting the results of accounting research and explaining and illustrating related research methodology. The scope of acceptable articles should embrace any research methodology and any accounting-related subject, as long as the articles meet the standards established for publication in the journal ... No special sections should be necessary. The primary, but not exclusive, audience should be—as it is now—academicians, graduate students, and others interested in accounting research.”

The primary criterion for publication in *The Accounting Review* is the significance of the contribution an article makes to the literature. Topical areas of interest to the journal include accounting information systems, auditing and assurance services, financial accounting, management accounting, taxation, and all other areas of accounting, broadly defined. The journal is also open to all rigorous research methods.

The efficiency and effectiveness of the editorial review process is critically dependent upon the actions of both authors submitting papers and the reviewers. Authors accept the responsibility of preparing research papers at a level suitable for evaluation by independent reviewers. Such preparation, therefore, should include subjecting the manuscript to critique by colleagues and others and revising it accordingly prior to submission. The review process is not to be used as a means of obtaining feedback at early stages of developing the research.

Reviewers and editors are responsible for providing constructive and prompt evaluations of submitted research papers based on the significance of their contribution and on the rigor of analysis and presentation.

MANUSCRIPT PREPARATION AND STYLE

The Accounting Review's manuscript preparation guidelines follow *The Chicago Manual of Style* (15th ed.; University of Chicago Press). Another helpful guide to usage and style is *The Elements of Style*, by William Strunk, Jr., and E. B. White (Macmillan). Spelling follows *Webster's Collegiate Dictionary*.

FORMAT

1. All manuscripts should be formatted in 12-point font on 8 1/2 x 11" paper and should be double-spaced, except for indented quotations.
2. Manuscripts should be as concise as the subject and research method permit, generally not to exceed 7,000 words.
3. Margins should be at least one inch from top, bottom, and sides.
4. To promote anonymous review, authors should not identify themselves directly or indirectly in their papers or in experimental test instruments included with the submission. Single authors should not use the editorial “we”.
5. A cover page should show the title of the paper, all authors' names, titles and affiliations, email addresses, and any acknowledgments.
6. The American Accounting Association encourages use of gender-neutral language in its publications.

Pagination: All pages, including tables, appendices and references, should be serially numbered. Major sections should be numbered in Roman numerals. Subsections should not be numbered.

Numbers: Spell out numbers from one to ten, except when used in tables and lists, and when used with mathematical, statistical, scientific, or technical units and quantities, such as distances, weights and measures. For example: *three days*; *3 kilometers*; *30 years*. All other numbers are expressed numerically.

Percentages and Decimal Fractions: In nontechnical copy use the word percent in the text.

Hyphens: Use a hyphen to join unit modifiers or to clarify usage. For example: *a cross-sectional equation*; *re-form*. See *Webster's* for correct usage.

Keywords: The abstract must be followed by at least three keywords to assist in indexing the paper and identifying qualified reviewers.

ABSTRACT/INTRODUCTION

An Abstract of about 100 words (150 maximum) should be presented on a separate page immediately preceding the text. The Abstract should concisely inform the reader of the manuscript's topic, its methods, and its findings. The Keywords statement should appear immediately below the Abstract. The text of the paper should start with a section labeled "I. Introduction," which provides more details about the paper's purpose, motivation, methodology, and findings. Both the Abstract and the Introduction should be relatively nontechnical, yet clear enough for an informed reader to understand the manuscript's contribution. The manuscript's title, but neither the author's name nor other identification designations, should appear on the Abstract page.

TABLES AND FIGURES

The author should note the following general requirements:

1. Each table and figure (graphic) should appear on a separate page and should be placed at the end of the text. Each should bear an Arabic number and a complete title indicating the exact contents of the table or figure. Tables and figures should define each variable. The titles and definitions should be sufficiently detailed to enable the reader to interpret the tables and figures without reference to the text.
2. A reference to each graphic should be made in the text.
3. The author should indicate where each graphic should be inserted in the text.
4. Graphics should be reasonably interpreted without reference to the text.
5. Source lines and notes should be included as necessary.

Equations: Equations should be numbered in parentheses flush with the right-hand margin.

DOCUMENTATION

Citations: Within-text citations are made using an author-year format. Cited works must correspond to the list of works listed in the "References" section. Authors should make an effort to include the relevant page numbers in the within-text citations.

1. In the text, works are cited as follows: author's last name and year, without comma, in parentheses. For example: one author, (Berry 2003); two authors, (Fehr and Schmidt 2003); three or more authors, (Dechow et al. 1998); more than one work cited, (Cole and Yakushiji 1984; Dechow et al. 1995; Levitt 1998); with two works by the same author(s), (Nelson 2003, 2005).
2. Unless confusion would result, do not use "p." or "pp." before page numbers. For example, (Dechow and Dichev 2002, 41–42).
3. When the reference list contains two or more works by the same author (as the only author or first of two or more authors) published in the same year, the suffix a, b, etc., is appended to the date in the within-text citations and in the "References" section. For example, (Johansson 2004a, 2004b, 2004c; Baiman and Rajan 2002a, 2002b; Dhaliwal et al. 2005a, 2005b).
4. When the author's name is mentioned in the text, it need not be repeated in the citation. For example: "Cohen et al. (2005) provide ..."
5. Citations to institutional works should use acronyms or short titles where practicable. For example: (NCFRR, The Treadway Commission 1987).
6. If the paper refers to statutes, legal treatises, or court cases, citations acceptable in law reviews should be used.

Reference List: Every manuscript must include a "References" section that contains only those works cited within the text. Each entry should contain all information necessary for unambiguous identification of the published work. Use the following formats (which follow *The Chicago Manual of Style*):

1. Arrange citations in *alphabetical order* according to the surname of the first author or the name of the institution or body responsible for the published work.
2. Use authors' initials instead of proper names.
3. For two or more authors, separate authors with a comma, including a comma before "and" (Dechow, P. M., R. Sloan, and A. Sweeney).
4. Date of publication follows the author's (authors') name(s).
5. Titles of journals or newspapers should not be abbreviated.

Sample entries are as follows:

- Baiman, S., and M. Rajan. 2002a. The role of information and opportunism in the choice of buyer-supplier relationships. *Journal of Accounting Research* 40 (2): 247–278.
- , and ———. 2002b. Incentive issues in inter-firm relationships. *Accounting, Organizations and Society* 27 (3): 213–238.
- Berry, R. 2003. Testimony before the Senate Committee on Homeland Security and Governmental Affairs Permanent Subcommittee on Investigations. November 18. Available at: <http://hsgac.senate.gov/files/111803berry.pdf>.
- Cohen, D., A. Dey, and T. Lys. 2005. The Sarbanes Oxley Act of 2002: Implications for compensation structure and risk-taking incentives of CEOs. Working paper, New York University, University of Chicago, and Northwestern University.
- Cole, R., and T. Yakushiji, eds. 1984. *The American and Japanese Auto Industries in Transition*. Ann Arbor, MI: University of Michigan.
- Dechow, P. M., R. Sloan, and A. Sweeney. 1995. Detecting earnings management. *The Accounting Review* 70 (2): 193–225.
- , S. P. Kothari, and R. L. Watts. 1998. The relation between earnings and cash flows. *Journal of Accounting and Economics* 25: 133–168.
- , and I. Dichev. 2002. The quality of accruals and earnings: The role of accrual estimation errors, *The Accounting Review* 77 (Supplement): 35–59.
- Dhaliwal, D., M. Erickson, and O. Li. 2005a. Shareholder income taxes and the relation between earnings and returns. *Contemporary Accounting Research* 22: 587–616.
- , L. Krull, O. Li, and W. Moser. 2005b. Dividend taxes and implied cost of equity capital. *Journal of Accounting Research* 43: 675–708.
- Easton, P. 2003. Discussion of: The predictive value of expenses excluded from pro forma earnings. *Review of Accounting Studies* 8: 175–183.
- Engel, E., R. Hayes, and X. Wang. 2007. The Sarbanes-Oxley Act and firms' going-private decisions. *Journal of Accounting and Economics* (forthcoming).
- Financial Accounting Standards Board (FASB). 2006. *Accounting for Uncertainty in Income Taxes, and Interpretation of FASB Statement No. 109*. FASB Interpretation No. 48. Financial Accounting Series. Norwalk, CT: FASB.
- Fehr, E., and K. Schmidt. 2003. A theory of fairness, competition, and cooperation. In *Advances in Behavioral Economics*, edited by C. Camerer, G. Loewenstein, and M. Rabin, 271–296. New York, NY: Princeton University Press.
- Janis, I. L. 1982. *Groupthink: Psychological Studies of Policy Decisions and Fiascoes*. Boston, MA: Houghton Mifflin.
- Levitt, A. 1998. The numbers game. Speech delivered at New York University, Center for Law and Business, September 28.
- Maggi, G. 1999. The value of commitment with imperfect observability and private information. *RAND Journal of Economics* (Winter) 30: 555–574.
- National Commission on Fraudulent Reporting (the Treadway Commission). 1987. *Report of the National Commission on Fraudulent Financial Reporting*. Washington, D.C.: NCFRR.
- Nelson, M. W. 2003. Behavioral evidence on the effects of principles- and rules-based standards. *Accounting Horizons* 17 (1): 91–104.
- . 2005. A review of experimental and archival conflicts-of-interest research in auditing. In *Conflicts of Interest: Challenges and Solutions in Business, Law, Medicine, and Public Policy*, edited by D. A. Moore, D. M. Cain, G. Loewenstein, and M. H. Bazerman. Cambridge, U.K.: Cambridge University Press.
- Rigdon, E. E., R. E. Schumacker, and W. Wothke. 1998. A comparative review of interaction and nonlinear modeling. In *Interaction and Nonlinear Effects in Structural Equation Modeling*, edited by R. E. Schumacker, and G. A. Marcoulides, 1–16. Mahwah, NJ: Erlbaum Associates.
- Schultz, E., and T. Francis. 2002. Companies profit on workers' deaths through "dead peasants" insurance. *Wall Street Journal* (April 19): 1.
- Securities and Exchange Commission (SEC). 2002. *Certification of Disclosure in Companies' Quarterly and Annual Reports*. Release Nos. 33-8124, 34-46427. Washington, D.C.: SEC.
- U.S. House of Representatives. 2002. The Sarbanes-Oxley Act of 2002. Public Law 107-204 [H. R. 3763]. Washington, D.C.: Government Printing Office.

Footnotes: Footnotes are not used for documentation. Textual footnotes should be used only for extensions and useful excursions of information that if included in the body of the text might disrupt its continuity. Footnotes should be inserted using the “footnote” or “endnote” feature of the word processing software which will automatically number the footnotes throughout the manuscript with superscript Arabic numerals.

SUBMISSION OF MANUSCRIPTS

Authors should note the following guidelines for submitting manuscripts:

1. Manuscripts currently under consideration by another journal or publisher should not be submitted. The author must state upon submission that the work is not submitted or published elsewhere.
2. For manuscripts reporting on field surveys or experiments: If the additional documentation (e.g., questionnaire, case, interview schedule) is sent as a separate file, then all information that might identify the author(s) must be deleted from the instrument.
3. Through May 31, 2011, manuscripts should be submitted via email as Microsoft Word or PDF file to the Senior Editor, Steven Kachelmeier, at email address tar@mcombs.utexas.edu. Please submit separate files for (1) the manuscript's title page with identifying information (not forwarded to reviewers), (2) the manuscript with title page and all other identifying information removed, and (3) any necessary supplemental files, such as experimental instructions and/or response memoranda on invited revisions. The body or subject line of the email should also indicate the submission fee receipt code from AAA (see step 5 below).
4. Beginning June 1, 2011, please check online at <https://aaahq.org/AAAforms/journals/tarsubmit.cfm> for instructions on how to submit a new manuscript to the incoming editor, John Harry Evans III. Revisions of manuscripts originally submitted before June 1, 2011 should be sent via email to the incoming editor, John Harry Evans III, at tar@katz.pitt.edu.
5. A nonrefundable submission fee is required unless explicitly waived by the editor for invited revisions of previous submissions. The submission fee in U.S. funds is \$200.00 for members and \$400.00 for non-members of the AAA payable by credit card (VISA or MasterCard only). The payment form is available online at: <https://aaahq.org/AAAforms/journals/tarsubmit.cfm>. If you are unable to pay by credit card or have any questions, please contact the AAA Member Services Team at (941) 921-7747 or info@aaahq.org.
6. Revisions must be submitted within 12 months from the decision letter inviting a revision.

COMMENTS

The Accounting Review welcomes submissions of comments on previous *TAR* articles. Comments on articles previously published in *The Accounting Review* will generally be reviewed by two reviewers, usually including an author of the original article (to assist the editor in evaluating whether the submitted comment represents the prior article accurately) and an additional reviewer who is independent of the original article. If a comment is accepted for publication, the original author will generally be invited to reply. All other editorial requirements, as enumerated above, apply to proposed comments.

POLICY ON REPRODUCTION

An objective of *The Accounting Review* is to promote the wide dissemination of the results of systematic scholarly inquiries into the broad field of accounting.

Permission is hereby granted to reproduce any of the contents of the *Review* for use in courses of instruction, as long as the source and American Accounting Association copyright are indicated in any such reproductions.

Written application must be made to the American Accounting Association, 5717 Bessie Drive, Sarasota, FL 34233-2399, for permission to reproduce any of the contents of the *Review* for use in other than courses of instruction—e.g., inclusion in books of readings or in any other publications intended for general distribution. In consideration for the grant of permission by the *Review* in such instances, the applicant must notify the author(s) in writing of the intended use to be made of each reproduction.

Except where otherwise noted in articles, the copyright interest has been transferred to the American Accounting Association. Where the author(s) has (have) not transferred the copyright to the Association, applicants must seek permission to reproduce (for all purposes) directly from the author(s).

POLICY ON DATA AVAILABILITY

The AAA's Executive Committee policy (originally adopted in 1989, and amended in 2009) is that the objective of the Association-wide journals (*The Accounting Review*, *Accounting Horizons*, *Issues in Accounting Education*) is to provide the widest possible dissemination of knowledge based on systematic scholarly inquiries into accounting as a field of professional research, and educational activity. To fulfill this objective, authors are encouraged to make their data available for use by others in extending or replicating results reported in their articles.

MAY 2011 PLACEMENT ADS

The deadline for free position ads to be included in this section is two months prior to the date of publication. *The Accounting Review* is published in January, March, May, July, September, and November. Position ads in the "Placement Ads" are free with the purchase of a job posting in the AAA Career Center. For more information on how to purchase a job posting in the Career Center and receive your free position ad in *The Accounting Review*, go to the AAA website at <http://aaahq.org> and click on "Career Center," or call our office at 941-921-7747.

- 301 THE UNIVERSITY OF GEORGIA, Department of Housing and Consumer Economics College of Family and Consumer Sciences is seeking candidates for an Endowed Professor in Family Financial Planning with an emphasis in estate planning, planned giving, or comprehensive financial planning. Nine-month academic appointment. An endowed professorship is associated with this position. Position begins August 1, 2011. Qualifications: Significant academic contributions to the field of financial planning. Demonstrated record of outstanding scholarly achievement in research productivity, external funding, and successful college teaching experience desired. Current academic rank should be that of Professor, or work should be commensurate with the rank of Professor. For a complete position description and more information, please visit our website at: <http://www.fr.uga.edu/hace-prof.pdf>. Application Procedure: To receive full consideration, applications should be submitted prior to April 3, 2011; however, the position will remain open until filled. Application materials should include a cover letter of interest, which outlines relevant experience, supports your fit for the position, as well as a current *curriculum vitae*. Email submissions with attachments are preferred. Please send to: executivesearch@uga.edu. Confidential requests for information should be directed to: Executive Michael Luthi, Search Consultant, Executive and Faculty Search Group, (706) 542-1837. The University of Georgia is an Equal Opportunity/Affirmative Action Institution.
- 302 THE UNIVERSITY OF ALABAMA, Culverhouse School of Accountancy invites applications for a tenure-track appointment at the Assistant Professor level beginning August 16, 2011. Candidates should have an earned doctorate (or expect to complete their doctorate by August 2011) and exhibit excellence in teaching and the capacity for research publishable in top-tier accounting journals. The faculty member can expect to teach undergraduate- and graduate-level accounting courses, although current needs are driven by increasing graduate enrollment. The faculty member will have the ability to work with doctoral students and engage in University, College, and Departmental service activities. We welcome applications in all areas of accounting research and teaching. Screening of candidates will begin immediately and will continue until the position is filled. Salary is competitive. Candidates selected for interview will be required to submit a disclosure and consent form authorizing a background investigation. Only online applications are accepted. Candidates must apply online at: <https://facultyjobs.ua.edu>. Required documents include: resume/*curriculum vitae*, dissertation proposal and/or example of scholarly work, and a listing of four references. The University of Alabama is an Equal Opportunity Educational Institution/Employer. Women and Minorities are encouraged to apply.

- 303 CENTRAL CONNECTICUT STATE UNIVERSITY, Accounting Department invites applications for a tenure-track Assistant/Associate Professor (#C11-058) beginning Fall 2011. Teaching interest in financial accounting and/or auditing is preferred. Candidates must have a commitment to excellence in teaching and scholarship. Qualifications: Earned doctorate in accounting or equivalent from an accredited school preferred, ABD with a 2011 expected completion date required. Professional certifications and/or experience are desirable. Teaching interest in financial accounting and/or auditing is preferred. Candidates must have a commitment to excellence in teaching and scholarship. To apply: Complete the Application Cover Page at: <http://www.ccsu.edu/AffAction/>. Send the cover page, along with a letter of application, your resume, and the names, addresses, and telephone numbers of three references via email to: Professor Lawrence Grasso, GrassoLa@ccsu.edu. Applications received by February 15, 2011 will receive preference of review. CCSU aggressively pursues a program of equal employment and educational opportunity and affirmative action. Members of all underrepresented groups, women, veterans, and persons with disabilities are invited and encouraged to apply.
- 304 UNIVERSITY OF SOUTH CAROLINA, School of Accounting is accepting applications for a tenure-track position beginning Fall 2011. Candidates should have primary teaching and research interests in international accounting, and must possess a doctorate or expect to complete it by August 2011. In addition to being committed to excellent teaching, candidates must be ambitious and motivated researchers who are committed to publishing in top-tier accounting journals. Salaries are competitive and commensurate with experience and achievements. All faculty searches are subject to the availability of funding. Submit *curriculum vitae*, working papers, and defended dissertation proposal with data to: Professor Al Leitch, Darla Moore School of Business, University of South Carolina, Columbia, SC 29208; or Email: leitch@moore.sc.edu. The University of South Carolina is an AA/EOE. Minorities and women are strongly encouraged to apply.
- 305 SAINT LOUIS UNIVERSITY, a Catholic, Jesuit institution dedicated to student learning, research, health care, and service, is seeking applicants for a tenure-track position in Accounting at the Senior Assistant/Associate/Full Professor level, beginning Fall 2011. Candidate must demonstrate a commitment to teaching excellence and quality research and publication. Service to the department and University are also required in addition to teaching and scholarship. Professional certification is desirable. Teaching interest is open and all areas will be considered except financial accounting. A Ph.D. in Accounting is required. Dissertation-stage candidates could be considered under exceptional circumstances. Compensation package is competitive. Interested candidates must apply online at <http://jobs.slu.edu> (Req. 20100775), and submit a cover letter and current *curriculum vitae*. All other correspondence regarding this position can be sent to: Dr. Ananth Seetharaman, Chair of Accounting, John Cook School of Business, Saint Louis University, 3674 Lindell Boulevard, St. Louis, MO 63108. Saint Louis University is an Affirmative Action/Equal Opportunity Employer, and encourages nominations of and application from women and minorities.
- 306 SABANCI UNIVERSITY (SU), School of Management invites applications for full-time/visiting positions in Accounting for the year 2011–2012. While pre-doctorate candidates are considered, the ideal candidate possesses a Ph.D. and a track record of excellence in research and teaching. SU, a private, innovative academic institution, offers excellent facilities and a light teaching load amid a modern campus located in Istanbul, Turkey. The University admits top-ranking faculty and students and offers undergraduate, M.B.A., Executive M.B.A. and Ph.D. programs in Operations Management, Organization Studies, and Finance. It provides faculty excellent support and competitive compensation, including housing. English is the medium of instruction. For further information, visit: <http://www.sabanciuniv.edu/ybf/eng/>. Applicants are requested to complete the online application form at: <https://crm.sabanciuniv.edu/acadev/application.php>, using ACC2011 as the Code Reference, and to send their cv and the names of three references to: Ulf Nilsson, nilsson@sabanciuniv.edu or Mine Aksu, maksu@sabanciuniv.edu. We will be interviewing at the Annual Meeting.

- 307 QUINNIPIAC UNIVERSITY invites applications for a full-time, tenure-track accounting faculty position, rank open, to start in August 2011. Consideration will be given to individuals interested in a full-time visiting position for the 2011–12 academic year. Primary teaching responsibilities will be in the financial accounting area. All candidates must hold a doctorate in accounting from an AACSB-accredited school (ABD considered for the visiting position). Successful candidates must demonstrate a strong potential for or have an established record of excellent teaching, published research, and significant service. Qualified candidates should apply online at <https://careers.quinnipiac.edu>. Application materials should include a *curriculum vitae*, evidence of teaching success, and the names of three references. Other inquiries should be directed to: Kathleen Simione, Chair of the Accounting Department at: accountingsearch@quinnipiac.edu. Quinnipiac University has a strong commitment to the principles and practices of diversity throughout the University community and welcomes candidates who would enhance that diversity.
- 308 OREGON STATE UNIVERSITY invites applicants for two anticipated full-time (1.0 FTE), nine-month, tenure-track positions in Accounting at the rank of Assistant Professor to begin September 2011. Teaching responsibilities include graduate and undergraduate accounting courses. Scholarly activities are expected to result in peer-reviewed journal publications. Required qualifications include an appropriate earned doctorate from an AACSB International accredited university or equivalent and evidence of potential for scholarly publication and teaching effectiveness. Preferred qualifications include professional certification and a demonstrable commitment to promoting and enhancing diversity. Information about the position and the University's online application process can be accessed at: <http://oregonstate.edu/jobs>, Posting No. 006857. Applications, including (1) a letter of interest clearly indicating interest in the position, (2) cv, (3) contact information for three professional references, (4) recent teaching evaluations and summary information of courses taught, and (5) evidence of scholarly work should be submitted by 3/15/2011. OSU is an AA/EOE.
- 309 TOWSON UNIVERSITY, College of Business and Economics invites applications for the position of Chair of the Department of Accounting, effective Summer or Fall 2011. The College of Business and Economics is an AACSB-accredited institution with more than 3,200 undergraduate business majors. The Department of Accounting has 18 full-time faculty members, more than 600 undergraduate majors, and more than 60 graduate students in a joint master's program in accounting with the University of Baltimore. Additional information can be found at: <http://www.new.towson.edu/cbe/>. Applicants should possess a Ph.D./D.B.A. with a concentration in accounting or closely related field from an AACSB-accredited university, and must provide evidence of effective teaching and intellectual contributions. Experience with the AACSB-accreditation process and requirements are highly desirable. This is a renewable 12-month administrative appointment at the Clinical Associate Professor level for an expected initial period of five years, with a reduced teaching load. Compensation is competitive. Review of applications will begin immediately. Interested candidates should submit a letter of application, current *curriculum vitae*, recent teaching evaluations (if applicable), and the names of three references to: Dr. Andrew Schiff, Search Committee Chair, Towson University, Department of Accounting, Stephens Hall, Room 102, 8000 York Road, Towson, MD 21252-0001, or Email: aschiff@towson.edu. Towson University is an Equal Opportunity/Affirmative Action Employer and has a strong institutional commitment to diversity. Women, minorities, persons with disabilities, and veterans are encouraged to apply.

- 310 WEST VIRGINIA UNIVERSITY, Department of Accounting and MIS in the College of Business and Economics invites applications for a tenure-track appointment at the Assistant or Associate Professor level beginning in August 2011. Our primary need is in Accounting Information Systems, but we welcome applications in all areas of accounting research and teaching. We anticipate filling at least one position. West Virginia University is a land-grant institution located in Morgantown, West Virginia, close to major metropolitan centers in the northeast. The Department of Accounting and MIS is separately accredited by the AACSB and offers undergraduate and graduate degrees, including a planned doctoral program. Applicants should have an earned doctorate in accounting (or equivalent) from an AACSB-accredited institution and exhibit effectiveness in teaching and the capacity for research publishable in respected scholarly outlets. CPA or other professional certification preferred. Interested candidates should submit a current *vita*, a research paper suitable for presentation at a workshop, the names and contact information of three references, and evidence of teaching ability to: Dr. Barbara Apostolou, Chair of the Search Committee, via email at: Barbara.Apostolou@mail.wvu.edu. Screening of candidates will begin immediately and continue until the position is filled. West Virginia University is an Equal Employment Opportunity/Affirmative Action Institution. The university does not discriminate on the basis of age, color, disability, national origin, race, religion, sexual orientation, or veteran status. West Virginia University is the recipient of an NSF ADVANCE award for gender equity. Minorities, persons with disabilities, females, and other protected class members are encouraged to apply. This offer of employment is conditioned upon approval by all the appropriate governmental authorities.
- 311 GEORGIA COLLEGE & STATE UNIVERSITY, J. Whitney Bunting College of Business is accepting applications for an academic-year tenure-track position in accounting beginning August 2011. The successful candidate will be able to teach introductory, upper-level, and graduate accounting courses. Preference will be given to candidates with a record of successful teaching in the area of financial accounting. Qualified candidates will have completed, or be near completion, of a Ph.D. or D.B.A. degree in Accounting from a business school with AACSB Accreditation. Professional experience and certification are considered very positive complements to academic qualifications. Excellence in teaching, scholarly activity, and university/community service are requirements for promotion and tenure. To be considered, applicants must apply electronically at: <http://www.gcsujobs.com>
- 312 ROBERT MORRIS UNIVERSITY (RMU), founded in 1921, has offered high-quality academic programs in business to students from the Pittsburgh region. RMU's School of Business, accredited by AACSB International, seeks qualified candidates with teaching interests in financial/manAGERIAL/cost accounting or accounting information systems. Qualified candidates should have an earned doctorate in Accounting from an AACSB-accredited school or be ABD near completion. Candidates must demonstrate excellence in teaching and research in their disciplines. Candidates should submit a letter of interest, current *vita*, and three academic and/or professional references to: Business@rmu.edu, c/o Dr. Jacobs, Dean, or mail to: Robert Morris University, 6001 University Boulevard, Moon Township, PA 15108, c/o Dr. Jacobs. Visit <http://www.rmu.jobs> for position descriptions. Anticipated start date is Fall 2011 and dependent upon budgetary approval. RMU is committed to increasing diversity in its community and actively pursues individuals from all backgrounds. EOE

- 313 PURDUE UNIVERSITY, Krannert Graduate School of Management invites applicants for one or more Continuous Term Lecturer appointment(s) in accounting beginning Fall 2011. Candidates will be expected to teach both Introduction to Tax and Advanced Taxation and be capable of teaching both financial and managerial accounting, primarily at the undergraduate level. A full-time teaching load is typically 24 credit hours per year. A successful candidate will have a demonstrated record of excellence in teaching preferably at comparable schools. A minimum of a master's degree with a professional certification in accounting is required; a Ph.D. with a concentration in accounting or related area is strongly preferred. Applications will be reviewed beginning February 15, 2011 and should include a *vitae*, evidence of teaching performance, and three letters of reference. Applications should be sent to: Professor Susan Watts, Krannert Graduate School of Management, Purdue University, 403 W. State Street, West Lafayette, IN 47907-2056. Purdue University is an Equal Opportunity/Equal Access/Affirmative Action Employer fully committed to achieving a diverse workforce.
- 314 UNIVERSITY OF WISCONSIN-STEVENS POINT is seeking candidates for an Assistant Professor of Accounting. The successful candidate will teach introductory and advanced courses in Accounting, with preference given to an interest in Auditing. Prefer Ph.D./D.B.A. in accounting from an AACSB-accredited institution. ABD will be considered. Ph.D. required for tenure. Send a letter of application including primary areas of teaching interest, resume or *vita*, copies of transcripts, and a list of three references (including phone numbers) to: Dr. Gary Mullins, Head School of Business and Economics, University of Wisconsin-Stevens Point, Stevens Point, WI 54481-3897; Email: gmullins@uwsp.edu. Screening will begin February 22, 2011. Position open until filled. Please see full position announcement at: <http://www.UWSP.edu>.
- 315 UNIVERSITY OF WASHINGTON BOTHELL, Business Program invites applications for a tenured opening in Accounting at the Professor or Associate Professor level. The appointment will become effective in the Fall 2011. Applicants must have a Ph.D. in Accounting and a commitment to excellence in scholarship and teaching. University of Washington faculty members engage in research, teaching, and service. Candidates should expect to teach at all levels of the undergraduate and graduate curriculum. They must be committed to working with diverse student and community populations. This position is contingent upon available funding. The successful candidate will have an established record of publications in high-quality academic journals and demonstrated excellence in teaching. In addition, he or she will play an important role in the continued development of the recently launched Business Program Accounting Option and new initiatives in the accounting area. This will include program and curriculum development, scheduling, mentoring junior faculty, and hiring part-time faculty. To apply: Please send a letter describing your qualifications and interest in the position as well as your *curriculum vitae* to: Dr. Steve Holland Chair, Accounting Search Committee, Business Program, Box 358533, University of Washington Bothell, 18115 Campus Way NE, Bothell, WA 98011-8246; Phone: (425) 352-5232; Fax: (425) 352-5277; Email: sholland@uwb.edu. For additional information, please see our website at: <http://www.uwb.edu/bus/>.
- 316 TEXAS CHRISTIAN UNIVERSITY, M. J. Neeley School of Business, Department of Accounting invites applications for a second Assistant Professor position to begin in August 2011. Candidates should hold a doctoral degree or expect to complete their degree prior to the start date. We seek candidates who are both excellent teachers and excellent researchers, consistent with TCU's Teacher-Scholar model. Applications should include a cover letter, *curriculum vitae*, and letters from at least three references. Applications may be submitted in either hardcopy or electronically, although electronic copies are preferred. Please send applications or nominations to: Ray Pfeiffer, Department Chair & Professor of Accounting, Texas Christian University, M. J. Neeley School of Business, TCU Box 298530, Fort Worth, TX 76129; Phone: (817) 257-6435; Email: r.pfeiffer@tcu.edu. Review of applications will continue until the position is filled. EEO/AA.

For focused results in accounting and tax research, start here.

While the world of accounting information continuously transitions, professionals, educators, and a growing number of accounting students look to reliable resources for researching accounting trends, rules, issues, and history pertaining to the future of accounting.

ProQuest Accounting & Tax, our premier accounting research database, makes research time more effective with centralized access to top accounting journals, reference reports, proceedings, dissertations, and more.

Researchers can instantly cross-search over 2,300 publications, with hundreds in ongoing full text, plus dissertations and working papers covering accounting and tax topics.

ABI/INFORM subscribers can enhance their business database library with ProQuest Accounting & Tax, which provides a vast array of additional full-text accounting journals, trade publications, dissertations, and working papers not included in ABI/INFORM.

ProQuest Accounting & Tax includes:

- Highly ranked global and scholarly journals, including all journals from the American Accounting Association
- Accounting, tax management and auditing trade publications from top publishers such as CCH, Aspen, Thomson Professional and Regulatory Services, Source Media, Tax Management Inc., and Euromoney
- AICPA publications
- International Federation of Accountants (IFAC) standards
- Tax law reviews from top institutions such as New York University, Stanford, Vanderbilt, and University of Virginia
- Tax law and financial law publications from noted sources such as the American Bar Association and American Law Institute
- Global accounting and tax journals from the UK, Canada, Australia, France, Spain, Ireland, and more
- Financial advising, governance, and risk management publications from Incisive Media, CFA Institute, and Association for Financial Counseling and Planning
- Accounting history from publications such as *The Accounting Historian's Journal*
- Over 1,500 accounting dissertations and more than 7,800 working papers from SSRN, plus select conference proceedings



ProQuest Accounting & Tax

Visit www.proquest.com/go/accounting today to learn more, and sign up for a free trial!



- 317 MISSISSIPPI STATE UNIVERSITY–MERIDIAN, Division of Business, is accepting applications for the position of Assistant Professor of Accounting. Teaching emphasis will be in managerial/financial accounting. This is a tenure-track position beginning August 16, 2011. Minimum qualifications: Doctorate in Accounting from an AACSB-accredited institution or Doctorate in Business Administration with an emphasis in Accounting from an AACSB-accredited institution with demonstrated excellence in the field of accountancy. Preferred qualifications: Professional certification and practical experience in accounting. Responsibilities: Teach a variety of undergraduate and graduate courses in accounting, publishing high-quality research, and advising and mentoring students. Review of applications will begin immediately and continue until position is filled. Interested applicants must complete the Personal Data Information Form online at: <http://www.jobs.msstate.edu>, and send: (1) letter of application; (2) *vitae*; (3) official transcripts; (4) three letters of recommendation; and (5) the name, address, and telephone number of three additional references to: Dr. Kevin Ennis, Search Committee Chair, Division of Business, MSU–Meridian, 1000 Highway 19 North, Meridian, MS 39307. MSU is an AA/EOE.

For focused results in accounting and tax research, start here.

While the world of accounting information continuously transitions, professionals, educators, and a growing number of accounting students look to reliable resources for researching accounting trends, rules, issues, and history pertaining to the future of accounting.

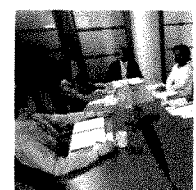
ProQuest Accounting & Tax, our premier accounting research database, makes research time more effective with centralized access to top accounting journals, reference reports, proceedings, dissertations, and more.

Researchers can instantly cross-search over 2,300 publications, with hundreds in ongoing full text, plus dissertations and working papers covering accounting and tax topics.

ABI/INFORM® subscribers can enhance their business database library with ProQuest Accounting & Tax, which provides a vast array of additional full-text accounting journals, trade publications, dissertations, and working papers not included in ABI/INFORM.

ProQuest Accounting & Tax includes:

- Highly ranked global and scholarly journals, including all journals from the American Accounting Association
- Accounting, tax management, and auditing trade publications from top publishers such as CCH, Aspen, Thomson Professional and Regulatory Services, Source Media, Tax Management Inc., and Euromoney
- AICPA publications
- International Federation of Accountants (IFAC) standards
- Tax law reviews from top institutions such as New York University, Stanford, Vanderbilt, and University of Virginia
- Tax law and financial law publications from noted sources such as the American Bar Association and American Law Institute
- Global accounting and tax journals from the UK, Canada, Australia, France, Spain, Ireland, and more
- Financial advising, governance, and risk management publications from Incisive Media, CFA Institute, and Association for Financial Counseling and Planning
- Accounting history from publications such as *The Accounting Historian's Journal*
- Over 1,500 accounting dissertations and more than 7,800 working papers from SSRN, plus select conference proceedings



ProQuest Accounting & Tax

Visit www.proquest.com/go/accounting
today to learn more, and sign up for a free trial!

